CGCTCAGGAT⬚⬚GACTTCG⬚⬚CGCT AG⬚⬚GATCGGATCCCGGC⬚⬚ ⬚⬚ATTATATAGCTCGATCGATC⬚
TTCTC⬚⬚TAT ⬚CCGCG⬚⬚⬚ATGG⬚⬚TATAT ACACAC⬚⬚⬚CG⬚CGCG⬚TAGC⬚TG⬚CTGATCT⬚
CCCC⬚⬚⬚⬚⬚⬚ACGC⬚⬚⬚⬚TCA⬚⬚⬚⬚CTTCGCATACGT⬚⬚⬚⬚⬚⬚CT⬚
CACAGAC⬚⬚⬚ACGC⬚⬚⬚⬚TCA⬚⬚⬚⬚CTTACTAACCAAT TCGCGAGCGGGUCGGGGATCGGGC GCGAG⬚

NCBI   Nucleotide

PubMed   Nucleotide   Protein   Genome   Structure   PMC   Taxonomy   OMIM   Books

Search   CoreNucleotide   ⬚ for   [Go] [Clear]

Limits   Preview/Index   History   Clipboard   Details

Display GenBank   ⬚ Show 5   ⬚ Send to   ⬚   Hide: ☐ sequence ☐ all but gene, CDS and mRNA features

Range: from begin   to end   ☐ Reverse complemented strand   Features: [+] [Refresh]

☐ **1:** AF206518. Reports West Nile virus i...[gi:7717200]   Links

Comment   Features   Sequence

```
LOCUS       AF206518                10975 bp    DNA     linear   VRL 08-MAY-2000
DEFINITION  West Nile virus isolate 2741, complete genome.
ACCESSION   AF206518
VERSION     AF206518.2  GI:7717200
KEYWORDS    .
SOURCE      West Nile virus
  ORGANISM  West Nile virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
            Flavivirus; Japanese encephalitis virus group.
REFERENCE   1  (bases 1 to 10975)
  AUTHORS   Anderson,J.F., Andreadis,T.G., Vossbrinck,C.R., Tirrell,S.,
            Wakem,E.M., French,R.A., Garmendia,A.E. and Van Kruiningen,H.J.
  TITLE     Isolation of West Nile virus from mosquitoes, crows, and a Cooper's
            hawk in Connecticut
  JOURNAL   Science 286 (5448), 2331-2333 (1999)
   PUBMED   10600741
REFERENCE   2  (bases 1 to 10975)
  AUTHORS   Vossbrinck,C.R., Anderson,J.F. and Andreadis,T.G.
  TITLE     Genome Sequence of West Nile Virus from Culex pipiens isolate
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 10975)
  AUTHORS   Anderson,J.F., Andreadis,T.G. and Vossbrinck,C.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-NOV-1999) Soil and Water, Connecticut Agricultural
            Experiment Station, 123 Huntington Street, New Haven, CT 06511, USA
REFERENCE   4  (bases 1 to 10975)
  AUTHORS   Anderson,J.F., Andreadis,T.G. and Vossbrinck,C.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-MAY-2000) Soil and Water, Connecticut Agricultural
            Experiment Station, 123 Huntington Street, New Haven, CT 06511, USA
  REMARK    Sequence update by submitter
COMMENT     On May 8, 2000 this sequence version replaced gi:6636507.
FEATURES             Location/Qualifiers
     source          1..10975
                     /organism="West Nile virus"
                     /mol_type="genomic DNA"
                     /strain="Connecticut 1999"
                     /isolate="2741"
                     /specific_host="Culex pipiens"
                     /db_xref="taxon:11082"
                     /country="USA:Connecticut, Greenwich-Stanford town line"
     CDS             79..10380
                     /codon_start=1
                     /product="polyprotein precursor"
                     /protein_id="AAF20205.2"
                     /db_xref="GI:7717201"
                     /translation="MSKKPGGPGKSRAVNMLKRGMPRVLSLIGLKRAMLSLIDGKGPI
                     RFVLALLAFFRFTAIAPTRAVLDRWRGVNKQTAMKHLLSFKKELGTLTSAINRRSSKQ
                     KKRGGKTGIAVMIGLIASVGAVTLSNFQGKVMMTVNATDVTDVITIPTAAGKNLCIVR
```

```
             AMDVGYMCDDTITYECPVLSAGNDPEDIDCWCTKSAVYVRYGRCTKTRHSRRSRRSLT
             VQTHGESTLANKKGAWMDSTKATRYLVKTESWILRNPGYALVAAVIGWMLGSNTMQRV
             VFVVLLLLVAPAYSFNCLGMSNRDFLEGVSGATWVDLVLEGDSCVTIMSKDKPTIDVK
             MMNMEAANLAEVRSYCYLATVSDLSTKAACPTMGEAHNDKRADPAFVCRQGVVDRGWG
             NGCGLFGKGSIDTCAKFACSTKAIGRTILKENIKYEVAIFVHGPTTVESHGNYSTQVG
             ATQAGRFSITPAAPSYTLKLGEYGEVTVDCEPRSGIDTNAYYVMTVGTKTFLVHREWF
             MDLNLPWSSAGSTVWRNRETLMEFEEPHATKQSVIALGSQEGALHQALAGAIPVEFSS
             NTVKLTSGHLKCRVKMEKLQLKGTTYGVCSKAFKFLGTPADTGHGTVVLELQYTGTDG
             PCKVPISSVASLNDLTPVGRLVTVNPFVSMATANAKVLIELEPPFGDSYIVVGRGEQQ
             INHHWHKSGSSIGKAFTTTLKGAQRLAALGDTAWDFGSVGGVFTSVGKAVHQVFGGAF
             RSLFGGMSWITQGLLGALLLWMGINARDRSIALTFLAVGGVLLFLSVNVHADTGCAID
             ISRQELRCGSGVFIHNDVEAWMDRYKYYPETPQGLAKIIQKAHKEGVCGLRSVSRLEH
             QMWEAVKDELNTLLKENGVDLSVVVEKQEGMYKSAPKRLTATTEKLEIGWKAWGKSIL
             FAPELANNTFVVDGPETKECPTQNRAWNSLEVEDFGFGLTSTRMFLKVRESNTTECDS
             KIIGTAVKNNLAIHSDLSYWIESRLNDTWKLERAVLGEVKSCTWPETHTLWGDGILES
             DLIIPVTLAGPRSNHNRRPGYKTQNQGPWDEGRVEIDFDYCPGTTVTLSESCGHRGPA
             TRTTTESGKLITDWCCRSCTLPPLRYQTDSGCWYGMEIRPQRHDEKTLVQSQVNAYNA
             DMIDPFQLGLLVVFLATQEVLRKRWTAKISMPAILIALLVLVFGGITYTDVLRYVILV
             GAAFAESNSGGDVVHLALMATFKIQPVFMVASFLKARWTNQENILLMLAAVFFQMAYH
             DARQILLWEIPDVLNSLAVAWMILRAITFTTTSNVVVPLLALLTPGLRCLNLDVYRIL
             LLMVGIGSLIREKRSAAAKKKGASLLCLALASTGLFNPMILAAGLIACDPNRKRGWPA
             TEVMTAVGLMFAIVGGLAELDIDSMAIPMTIAGLMFAAFVISGKSTDMWIERTADISW
             ESDAEITGSSERVDVRLDDDGNFQLMNDPGAPWKIWMLRMVCLAISAYTPWAILPSVV
             GFWITLQYTKRGGVLWDTPSPKEYKKGDTTTGVYRIMTRGLLGSYQAGAGVMVEGVFH
             TLWHTTKGAALMSGEGRLDPYWGSVKEDRLCYGGPWKLQHKWNGQDEVQMIVVEPGKN
             VKNVQTKPGVFKTPEGEIGAVTLDFPTGTSGSPIVDKNGDVIGLYGNGVIMPNGSYIS
             AIVQGERMDEPIPAGFEPEMLRKKQITVLDLHPGAGKTRRILPQIIKEAINRRLRTAV
             LAPTRVVAAEMAEALRGLPIRYQTSAVPREHNGNEIVDVMCHATLTHRLMSPHRVPNY
             NLFVMDEAHFTDPASIAARGYISTKVELGEAAAIFMTATPPGTSDPFPESNSPISDLQ
             TEIPDRAWNSGYEWITEYTGKTVWFVPSVKMGNEIALCLQRAGKKVVQLNRKSYETEY
             PKCKNDDWDFVITTDISEMGANFKASRVIDSRKSVKPTIITEGEARVILGEPSAVTAA
             SAAQRRGRIGRNPSQVGDEYCYGGHTNEDDSNFAHWTEARIMLDNINMPNGLIAQFYQ
             PEREKVYTMDGEYRLRGEERKNFLELLRTADLPVWLAYKVAAAGVSYHDRRWCFDGPR
             TNTTILEDNNEVEVITKLGERKILRPRWIDARVYSDHQALKAFKDFASGKRSQIGLIEV
             LGKMPEHFMGKTWEALDTMYVVATAEKGGRAHRMALEELPDALQTIALIALLSVMTMG
             VFFLLMQRKGIGKIGLGGAVLGVATFFCWMAEVPGTKIAGMLLLSLLLMIVLIPEPEK
             QRSQTDNQLAVFLICVMTLVSAVAANEMGWLDKTKSDISSLFGQRIEVKENFSMGEFL
             LDLRPATAWSLYAVTTAVLTPLLKHLITSDYINTSLTSINVQASALFTLARGFPFVDV
             GVSALLLAAGCWGQVTLTVTVTAATLLFCHYAYMVPGWQAEAMRSAQRRTAAGIMKNA
             VVDGIVATDVPELERTTPIMQKKVGQIMLILVSLAAVVVNPSVKTVREAGILITAAAV
             TLWENGASSVWNATTAIGLCHIMRGGWLSCLSITWTLIKNMEKPGLKRGGAKGRTLGE
             VWKERLNQMTKEEFTRYRKEAIIEVDRSAAKHARKEGNVTGGHPVSRGTAKLRWLVER
             RFLEPVGKVIDLGCGRGGWCYYMATQKRVQEVRGYTKGGPGHEEPQLVQSYGWNIVTM
             KSGVDVFYRPSECCDTLLCDIGESSSSAEVEEHRTIRVLEMVEDWLHRGPREFCVKVL
             CPYMPKVIEKMELLQRRYGGGLVRNPLSRNSTHEMYWVSRASGNVVHSVNMTSQVLLG
             RMEKRTWKGPQYEEDVNLGSGTRAVGKPLLNSDTSKIKNRIERLRREYSSTWHHDENH
             PYRTWNYHGSYDVKPTGSASSLVNGVVRLLSKPWDTITNVTTMAMTDTTPFGQQRVFK
             EKVDTKAPEPPEGVKYVLNETTNWLWAFLAREKRPRMCSREEFIRKVNSNAALGAMFE
             EQNQWRSAREAVEDPKFWEMVDEEREAHLRGECHTCIYNMMGKREKKPGEFGKAKGSR
             AIWFMWLGARFLEFEALGFLNEDHWLGRKNSGGGVEGLGLQKLGYILREVGTRPGGKI
             YADDTAGWDTRITRADLENEAKVLELLDGEHRRLARAIIELTYRHKVVKVMRPAADGR
             TVMDVISREDQRGSGQVVTYALNTFTNLAVQLVRMMEGEGVIGPDDVEKLTKGKGPKV
             RTWLFENGEERLSRMAVSGDDCVVKPLDDRFATSLHFLNAMSKVRKDIQEWKPSTGWY
             DWQQVPFCSNHFTELIMKDGRTLVVPCRGQDELVGRARISPGAGWNVRDTACLAKSYA
             QMWLLLYFHRRDLRLMANAICSAVPVNWVPTGRTTWSIHAGGEWMTTEDMLEVWNRVW
             IEENEWMEDKTPVEKWSDVPYSGKREDIWCGSLIGTRARATWAENIQVAINQVRAIIG
             DEKYVDYMSSLKRYEDTTLVEDTVL"
ORIGIN
        1 gctgacaaac ttagtagtgt ttgtgaggat taacaacaat taacacagtg cgagctgttt
       61 cttagcacga agatctcgat gtctaagaaa ccaggagggc ccggcaagag ccgggctgtc
      121 aatatgctaa aacgcggaat gccccgcgtg ttgtccttga ttggactgaa gagggctatg
      181 ttgagcctga tcgacggcaa ggggccaata cgatttgtgt tggctctctt ggcgttcttc
      241 aggttcacag caattgctcc gacccgagca gtgctggatc gatggagagg tgtgaacaaa
      301 caaacagcga tgaaacacct tctgagtttt aagaaggaac tagggacctt gaccagtgct
      361 atcaatcggc ggagctcaaa acaaaagaaa agaggaggaa agaccggaat tgcagtcatg
      421 attggcctga tcgccagcgt aggagcagtt accctctcta acttccaagg gaaggtgatg
```

```
 481 atgacggtaa atgctactga cgtcacagat gtcatcacga ttccaacagc tgctggaaag
 541 aacctatgca ttgtcagagc aatggatgtg ggatacatgt gcgatgatac tatcacttat
 601 gaatgcccag tgctgtcggc tggtaatgat ccagaagaca tcgactgttg gtgcacaaag
 661 tcagcagttt acgtcaggta tggaagatgc accaagacac gccactcaag acgcagtcgg
 721 aggtcactga cagtgcagac acacggagaa agcactctag cgaacaagaa gggggcttgg
 781 atggacagca ccaaggccac aaggtacttg gtaaaaacag aatcatggat cttgaggaac
 841 cctggatatg ccctggtggc agccgtcatt ggttggatgc ttgggagcaa caccatgcag
 901 agagttgtgt ttgtcgtgct attgcttttg gtggccccag cttacagctt caactgcctt
 961 ggaatgagca acagagactt cttggaagga gtgtctggag caacatgggt ggatttggtt
1021 ctcgaaggcg acagctgcgt gactatcatg tctaaggaca agcctaccat cgatgtgaag
1081 atgatgaata tggaggcggc caacctggca gaggtccgca gttattgcta tttggctacc
1141 gtcagcgatc tctccaccaa agctgcgtgc ccgaccatgg gagaagctca caatgacaaa
1201 cgtgctgacc cagcttttgt gtgcagacaa ggagtggtgg acaggggctg gggcaacggc
1261 tgcggactat ttggcaaagg aagcattgac acatgcgcca aatttgcctg ctctaccaag
1321 gcaataggaa gaaccatctt gaaagagaat atcaagtacg aagtggccat ttttgtccat
1381 ggaccaacta ctgtggagtc gcacggaaac tactccacac aggttggagc cactcaggca
1441 gggagattca gcatcactcc tgcagcgcct tcatacacac taaagcttgg agaatatgga
1501 gaggtgacag tggactgtga accacggtca gggattgaca ccaatgcata ctacgtgatg
1561 actgttggaa caaagacgtt cttggtccat cgtgagtggt tcatggacct caacctccct
1621 tggagcagtg ctggaagtac tgtgtggagg aacagagaga cgttaatgga gtttgaggaa
1681 ccacacgcca cgaagcagtc tgtgatagca ttgggctcac aagagggagc tctgcatcaa
1741 gctttggctg gagccattcc tgtggaattt tcaagcaaca ctgtcaagtt gacgtcgggt
1801 catttgaagt gtagagtgaa gatggaaaaa ttgcagttga agggaacaac ctatggcgtc
1861 tgttcaaagg ctttcaagtt tcttgggact cccgcagaca caggtcacgg cactgtggtg
1921 ttggaattgc agtacactgg cacggatgga ccttgcaaag ttcctatctc gtcagtggct
1981 tcattgaacg acctaacgcc agtgggcaga ttggtcactg tcaacccttt tgtttcaatg
2041 gccacggcca acgctaaggt cctgattgaa ttggaaccac cctttggaga ctcatacata
2101 gtggtgggca gaggagaaca acagatcaat caccattggc acaagtctgg aagcagcatt
2161 ggcaaagcct ttacaaccac cctcaaagga gcgcagagac tagccgctct aggagacaca
2221 gcttgggact ttggatcagt tggaggggtg ttcacctcag ttgggaaggc tgtccatcaa
2281 gtgttcggag gagcattccg ctcactgttc ggaggcatgt cctggataac gcaaggattg
2341 ctgggggctc tcctgttgtg gatgggcatc aatgctcgtg ataggtccat agctctcacg
2401 tttctcgcag ttggaggagt tctgctcttc ctctccgtga acgtgcacgc tgacactggg
2461 tgtgccatag acatcagccg gcaagagctg agatgtggaa gtggagtgtt catacacaat
2521 gatgtggagg cttggatgga ccggtacaag tattaccctg aaacgccaca aggcctagcc
2581 aagatcattc agaaagctca taggaagga gtgtgcggtc tacgatcagt ttccagactg
2641 gagcatcaaa tgtgggaagc agtgaaggac gagctgaaca ctcttttgaa ggagaatggt
2701 gtggacctta gtgtcgtggt tgagaaacag gaggaatgt acaagtcagc acctaaacgc
2761 ctcaccgcca ccacggaaaa attggaaatt ggctggaagg cctggggaaa gagtatttta
2821 tttgcaccag aactcgccaa caacaccttt gtggttgatg gtccggagac caaggaatgt
2881 ccgactcaga atcgcgcttg gaatagctta gaagtggagg attttggatt tggtctcacc
2941 agcactcgga tgttcctgaa ggtcagagag agcaacacaa ctgaatgtga ctcgaagatc
3001 attggaacgg ctgtcaagaa caacttggcg atccacagtg acctgtccta ttggattgaa
3061 agcaggctca atgatacgtg gaagcttgaa agggcagttc tgggtgaagt caaatcatgt
3121 acgtggcctg agacgcatac cttgtggggc gatggaatcc ttgagagtga cttgataata
3181 ccagtcacac tggcgggacc acgaagcaat cacaatcgga gacctgggta caagacacaa
3241 aaccagggcc catgggacga aggccgggta gagattgact tcgattactg cccaggaact
3301 acggtcaccc tgagtgagag ctgcggacac cgtggacctg ccactcgcac caccacagag
3361 agcggaaagt tgataacaga ttggtgctgc aggagctgca ccttaccacc actgcgctac
3421 caaactgaca gcggctgttg gtatggtatg gagatcagac cacagagaca tgatgaaaag
3481 accctcgtgc agtcacaagt gaatgcttat aatgctgata tgattgaccc tttttcagttg
3541 ggccttctgg tcgtgttctt ggccacccag gaggtccttc gcaagaggtg gacagccaag
3601 atcagcatgc cagctatact gattgctctg ctagtcctgg tgtttgggg cattacttac
3661 actgatgtgt tacgctatgt catcttggtg ggggcagctc tcgcagaatc taattcggga
3721 ggagacgtgg tacacttggc gctcatggcg accttcaaga tacaaccagt gtttatggtg
3781 gcatcgtttc ttaaagcgag atggaccaac caggagaaca ttttgttgat gttggcggct
3841 gtttttcttt aaaatggctta tcacgatgcc cgccaaattc tgctctggga gatccctgat
3901 gtgttgaatt cactggcggt agcttggatg atactgagag ccataacatt cacaacgaca
3961 tcaaacgtgg ttgttccgct gctagccctg ctaacacccg ggctgagatg cttgaatctg
4021 gatgtgtaca ggatactgct gttgatggtc ggaataggca gcttgatcag ggagaagagg
4081 agtgcagctg caaaaaagaa aggagcaagt ctgctatgct tggctctagc ctcaacagga
4141 ctttttcaacc ccatgatcct tgctgctgga ctgattgcat gtgatcccaa ccgtaaacgc
4201 ggatggcccg caactgaagt gatgacagct gtcggcctaa tgtttgccat cgtcggaggg
4261 ctggcagagc ttgacattga ctccatggcc attccaatga ctatcgcggg gctcatgttt
4321 gctgctttcg tgatttctgg gaaatcaaca gatatgtgga ttgagagaac ggcggacatt
4381 tcctgggaaa gtgatgcaga aattacaggc tcgagcgaaa gagttgatgt gcggcttgat
```

```
4441 gatgatggaa acttccagct catgaatgat ccaggagcac cttggaagat atggatgctc
4501 agaatggtct gtctcgcgat tagtgcgtac accccctggg caatcttgcc ctcagtagtt
4561 ggattttgga taactctcca atacacaaag agaggaggcg tgttgtggga cactccctca
4621 ccaaaggagt acaaaaaggg ggacacgacc accggcgtct acaggatcat gactcgtggg
4681 ctgctcggca gttatcaagc aggagcgggc gtgatggttg aaggtgtttt ccacaccctt
4741 tggcatacaa caaaggagc cgctttgatg agcggagagg gccgcctgga cccatactgg
4801 ggcagtgtca aggaggatcg actttgttac ggaggaccct ggaaattgca gcacaagtgg
4861 aacgggcagg atgaggtgca gatgattgtg gtggaacctg gcaagaacgt taagaacgtc
4921 cagacgaaac caggggtgtt caaaacacct gaaggagaaa tcggggccgt gactttggac
4981 ttccccactg gaacatcagg ctcaccaata gtggacaaaa acggtgatgt gattgggctt
5041 tatggcaatg gagtcataat gcccaacggc tcatacataa gcgcgatagt gcagggtgaa
5101 aggatggatg agccaatccc agccggattc gaacctgaga tgctgaggaa aaaacagatc
5161 actgtactgg atctccatcc cggcgccggt aaaacaagga ggattctgcc acagatcatc
5221 aaagaggcca taaacagaag actgagaaca gccgtgctag caccaaccag ggttgtggct
5281 gctgagatgg ctgaagcact gagaggactg cccatccggt accagacatc cgcagtgccc
5341 agagaacata atggaaatga gattgttgat gtcatgtgtc atgctaccct cacccacagg
5401 ctgatgtctc ctcacagggt gccgaactac aacctgttcg tgatggatga ggctcatttc
5461 accgacccag ctagcattgc agcaagaggt tacatttcca caaaggtcga gctaggggag
5521 gcggcggcaa tattcatgac agccaccca ccaggcactt cagatccatt cccagagtcc
5581 aattcaccaa tttccgactt acagactgag atcccggatc gagcttggaa ctctggatac
5641 gaatggatca cagaatacac cgggaagacg gtttggtttg tgcctagtgt caagatgggg
5701 aatgagattg ccctttgcct acaacgtgct ggaaagaaag tagtccaatt gaacagaaag
5761 tcgtacgaga cggagtaccc aaaatgtaag aacgatgatt gggactttgt tatcacaaca
5821 gacatatctg aaatgggggc taacttcaag gcgagcaggg tgattgacag ccggaagagt
5881 gtgaaaccaa ccatcataac agaaggagaa gcgagagtga tcctgggaga accatctgca
5941 gtgacagcag ctagtgccgc ccagagacgt ggacgtatcg gtagaaatcc gtcgcaagtt
6001 ggtgatgagt actgttatgg ggggcacacg aatgaagacg actcgaactt cgcccattgg
6061 actgaggcac gaatcatgct ggacaacatc aacatgccaa acggactgat cgctcaattc
6121 taccaaccag agcgtgagaa ggtatatacc atggatgggg aataccggct cagaggagaa
6181 gagagaaaaa actttctgga actgttgagg actgcagatc tgccagtttg gctggcttac
6241 aaggttgcag cggctggagt gtcataccac gaccggaggg ggtgctttga tggtcctagg
6301 acaaacacaa ttttagaaga caacaacgaa gtggaagtca tcacgaagct tggtgaaagg
6361 aagattctga ggccgcgctg gattgacgcc agggtgtact cggatcacca ggcactaaag
6421 gcgttcaagg acttcgcctc gggaaaacgt tctcagatag ggctcattga ggttctggga
6481 aagatgcctg agcacttcat ggggaagaca tgggaagcac ttgacaccat gtacgttgtg
6541 gccactgcag agaaaggagg aagagctcac agaatggccc tggaggaact gccagatgct
6601 cttcagacaa ttgccttgat tgccttattg agtgtgatga ccatgggagt attcttcctc
6661 ctcatgcagc ggaagggcat tggaaagata ggtttgggag gcgctgtctt gggagtcgcg
6721 accttttct gttggatggc tgaagttcca ggaacgaaga tcgccggaat gttgctgctc
6781 tcccttctct tgatgattgt gctaattcct gagccagaag agcaacgttc gcagacagac
6841 aaccagctag ccgtgttcct gatttgtgtc atgacccttg tgagcgcagt ggcagccaac
6901 gagatggggt tggctagataa gaccaagagt gacataagca gtttgtttgg gcaaagaatt
6961 gaggtcaagg agaatttcag catgggagag tttcttctgg acttgaggcc ggcaacagcc
7021 tggtcactgt acgctgtgac aacagcggtc ctcactccac tgctaaagca tttgatcacg
7081 tcagattaca tcaacaccct cattgacctca ataaacgttc aggcaagtgc actattcaca
7141 ctcgcgcgag gcttccccctt cgtcgatgtt ggagtgtcgg ctctcctgct agcagccgga
7201 tgctgggggac aagtcacccct caccgttacg gtaacagcgg caacactcct tttttgccac
7261 tatgcctaca tggttcccgg ttggcaagct gaggcaatgc gctcagccca gcggcggaca
7321 gcggccggaa tcatgaagaa cgctgtagtg gatggcatcg tggccacgga cgtcccagaa
7381 ttagagcgca ccacacccat catgcagaag aaagttggac agatcatgct gatcttggtg
7441 tctctagctg cagtagtagt gaacccgtct gtgaagacag tacgagaagc cggaattttg
7501 atcacggccg cagcggtgac gctttgggag aatggagcaa gctctgtttg gaacgcaaca
7561 actgccatcg gactctgcca catcatgcgt gggggttggt tgtcatgtct atccataaca
7621 tggacactca taaagaacat ggaaaaacca ggactaaaaa gaggtg999gc aaaaggacgc
7681 accttgggag aggtttggaa agaaagactc aaccagatga caaagaaga gttcactagg
7741 taccgcaaag aggccatcat cgaagtcgat cgctcagcgg caaaacacgc caggaaagaa
7801 ggcaatgtca ctggagggca tccagtctct aggggcacag caaaactgag atggctggtc
7861 gaacggaggt ttctcgaacc ggtcggaaaa gtgattgacc ttggatgtgg aagaggcggt
7921 tggtgttact atatggcaac ccaaaaaaga gtccaagaag tcagagggta cacaaagggc
7981 ggtcccggac atgaagagcc ccaactagtg caaagttatg gatggaacat tgtcaccatg
8041 aagagtggag tggatgtgtt ctacagacct tctgagtgtt gtgacaccct cctttgtgac
8101 atcggagagt cctcgtcaag tgctgaggtt gaagagcata ggacgattcg ggtccttgaa
8161 atggttgagg actggctgca ccgagggcca agggaatttt gcgtgaaggt gctctgcccc
8221 tacatgccga aagtcataga gaagatggag ctgctccaac gccggtatgg ggggggactg
8281 gtcagaaacc cactctcacg gaattccacg cacgagatgt attgggtgag tcgagcttca
8341 ggcaatgtgg tacattcagt gaatatgacc agccaggtgc tcctaggaag aatggaaaaa
```

```
 8401 aggacctgga agggaccccca atacgaggaa gacgtaaact tgggaagtgg aaccagggcg
 8461 gtgggaaaac ccctgctcaa ctcagacacc agtaaaatca agaacaggat tgaacgactc
 8521 aggcgtgagt acagttcgac gtggcaccac gatgagaacc acccatatag aacctggaac
 8581 tatcacggca gttatgatgt gaagcccaca ggctccgcca gttcgctggt caatggagtg
 8641 gtcaggctcc tctcaaaacc atgggacacc atcacgaatg ttaccaccat ggccatgact
 8701 gacactactc ccttcgggca gcagcgagtg ttcaaagaga aggtggacac gaaagctcct
 8761 gaaccgccag aaggagtgaa gtacgtgctc aacgagacca ccaactggtt gtgggcgttt
 8821 ttggccagag aaaaacgtcc cagaatgtgc tctcgagagg aattcataag aaaggtcaac
 8881 agcaatgcag ctttgggtgc catgtttgaa gagcagaatc aatggaggag cgccagagaa
 8941 gcagttgaag atccaaaatt ttgggagatg gtggatgagg agcgcgaggc acatctgcgg
 9001 ggggaatgtc acacttgcat ttacaacatg atgggaaaga gagagaaaaa acccggagag
 9061 ttcggaaagg ccaagggaag cagagccatt tggttcatgt ggctcggagc tcgctttctg
 9121 gagttcgagg ctctgggttt tctcaatgaa gaccactggc ttggaagaaa gaactcagga
 9181 ggaggtgtcg agggcttggg cctccaaaaa ctgggttaca tcctgcgtga agttggcacc
 9241 cggcctgggg gcaagatcta tgctgatgac acagctggct gggacacccg catcacgaga
 9301 gctgacttgg aaaatgaagc taaggtgctt gagctgcttg atggggaaca tcggcgtctt
 9361 gccagggcca tcattgagct cacctatcgt cacaaagttg tgaaagtgat gcgcccggct
 9421 gctgatggaa gaaccgtcat ggatgttatc tccagagaag atcagagggg gagtggacaa
 9481 gttgtcacct acgccctaaa cactttcacc aacctggccg tccagctggt gaggatgatg
 9541 gaaggggaag gagtgattgg cccagatgat gtggagaaac tcacaaaagg gaaaggaccc
 9601 aaagtcagga cctggctgtt tgagaatggg gaagaaagac tcagccgcat ggctgtcagt
 9661 ggagatgact gtgtggtaaa gcccctggac gatcgctttg ccacctcgct ccacttcctc
 9721 aatgctatgt caaaggttcg caaagacatc caagagtgga aaccgtcaac tggatggtat
 9781 gattggcagc aggttccatt ttgctcaaac catttcactg aattgatcat gaaagatgga
 9841 agaacactgg tggttccatg ccgaggacag gatgaattgg taggcagagc tcgcatatct
 9901 ccaggggccg gatggaacgt ccgcgacact gcttgtctgg ctaagtctta tgcccagatg
 9961 tggctgcttc tgtacttcca cagaagagac ctgcggctca tggccaacgc catttgctcc
10021 gctgtccctg tgaattgggt ccctaccgga agaaccacgt ggtccatcca tgcaggagga
10081 gagtggatga caacagagga catgttggag gtctggaacc gtgtttggat agaggagaat
10141 gaatggatgg aagacaaaac cccagtggag aaatggagtg acgtcccata ttcaggaaaa
10201 cgagaggaca tctggtgtgg cagcctgatt ggcacaagag cccgagccac gtgggcagaa
10261 aacatccagg tggctatcaa ccaagtcaga gcaatcatcg agatgagaa gtatgtggat
10321 tacatgagtt cactaaagag atatgaagac acaactttgg ttgaggacac agtactgtag
10381 atatttaatt aattgtaaat agacaatata agtatgcata aaagtgtagt tttatagtag
10441 tatttagtgg tgttagtgta aatagttaag aaaattttga ggagaaagtc aggccgggaa
10501 gttcccgcca ccggaagttg agtagacggt gctgcctgcg actcaacccc aggaggactg
10561 ggtgaacaaa gccgcgaagt gatccatgta agccctcaga accgtctcgg aaggaggacc
10621 ccacatgttg taacttcaaa gcccaatgtc agaccacgct acggcgtgct actctgcgga
10681 gagtgcagtc tgcgatagtg ccccaggagg actgggttaa caaaggcaaa ccaacgcccc
10741 acgcggccct agccccggta atggtgttaa ccaggcgaa aggactagag gttagaggag
10801 accccgcggt ttaaagtgca cggcccagcc tggctgaagc tgtaggtcag gggaaggact
10861 agaggttagt ggagaccccg tgccacaaaa caccacaaca aaacagcata ttgacacctg
10921 ggatagacta ggagatcttc tgctctgcac aaccagccac acggcacagt gcgcc
//
```

Apr 17 2007 11:10:07

Search CoreNucleotide ▓ for [   ] | Go | | Clear |

Limits   Preview/Index   History   Clipboard   Details

Display GenBank ▓ Show 5 ▓ Send to ▓   Hide: ☐ sequence ☐ all but gene, CDS and mRNA features

Range: from begin   to end   ☐ Reverse complemented strand   Features: |+| | Refresh |

☐ **1:** AF260967. Reports West Nile virus s...[gi:9930133]   Links

Features   Sequence

```
LOCUS       AF260967                11029 bp    RNA     linear   VRL 27-AUG-2000
DEFINITION  West Nile virus strain NY99-eqhs, complete genome.
ACCESSION   AF260967
VERSION     AF260967.1  GI:9930133
KEYWORDS    .
SOURCE      West Nile virus
  ORGANISM  West Nile virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
            Flavivirus; Japanese encephalitis virus group.
REFERENCE   1  (bases 1 to 11029)
  AUTHORS   Bowen,M., Meyer,R.F., McKinney,N., Morrill,W. and Lanciotti,R.
  TITLE     Complete genomic sequence of West Nile virus equine isolate New
            York 1999
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 11029)
  AUTHORS   Bowen,M., Meyer,R.F., McKinney,N., Morrill,W. and Lanciotti,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-APR-2000) Arbovirus Diseases Branch, Centers for
            Disease Control & Prevention, Rampart Road, Fort Collins, CO 80521,
            USA
FEATURES             Location/Qualifiers
     source          1..11029
                     /organism="West Nile virus"
                     /mol_type="genomic RNA"
                     /strain="NY99-eqhs"
                     /specific_host="horse"
                     /db_xref="taxon:11082"
                     /country="USA: New York"
                     /note="1999"
     CDS             97..10398
                     /codon_start=1
                     /product="polyprotein precursor"
                     /protein_id="AAG02038.1"
                     /db_xref="GI:9930134"
                     /translation="MSKKPGGPGKSRAVNMLKRGMPRVLSLIGLKRAMLSLIDGKGPI
                     RFVLALLAFFRFTAIAPTRAVLDRWRGVNKQTAMKHLLSFKKELGTLTSAINRRSSKQ
                     KKRGGKTGIAVMIGLIASVGAVTLSNFQGKVMMTVNATDVTDVITIPTAAGKNLCIVR
                     AMDVGYMCDDTITYECPVLSAGNDPEDIDCWCTKSAVYVRYGRCTKTRHSRRSRRSLT
                     VQTHGESTLANKKGAWMDSTKATRYLVKTESWILRNPGYALVAAVIGWMLGSNTMQRV
                     VFVVLLLLVAPAYSFNCLGMSNRDFLEGVSGATWVDLVLEGDSCVTIMSKDKPTIDVK
                     MMNMEAANLAEVRSYCYLATVSDLSTKAACPTMGEAHNDKRADPAFVCRQGVVDRGWG
                     NGCGLFGKGSIDTCAKFACSTKAIGRTILKENIKYEVAIFVHGPTTVESHGNYSTQVG
                     ATQAGRFSITPAAPSYTLKLGEYGEVTVDCEPRSGIDTNAYYVMTVGTKTFLVHREWF
                     MDLNLPWSSAGSTVWRNRETLMEFEEPHATKQSVIALGSQEGALHQALAGAIPVEFSS
                     NTVKLTSGHLKCRVKMEKLQLKGTTYGVCSKAFKFLGTPADTGHGTVVLELQYTGTDG
                     PCKVPISSVASLNDLTPVGRLVTVNPFVSVATANAKVLIELEPPFGDSYIVVGRGEQQ
                     INHHWHKSGSSIGKAFTTTLKGAQRLAALGDTAWDFGSVGGVFTSVGKAVHQVFGGAF
                     RSLFGGMSWITQGLLGALLLWMGINARDRSIALTFLAVGGVLLFLSVNVHADTGCAID
                     ISRQELRCGSGVFIHNDVEAWMDRYKYYPETPQGLAKIIQKAHKEGVCGLRSVSRLEH
```

```
                    QMWEAVKDELNTLLKENGVDLSVVVEKQEGMYKSAPKRLTATTEKLEIGWKAWGKSIL
                    FAPELANNTFVVDGPETKECPTQNRAWNSLEVEDFGFGLTSTRMFLKVRESNTTECDS
                    KIIGTAVKNNLAIHSDLSYWIESRLNDTWKLERAVLGEVKSCTWPETHTLWGDGILES
                    DLIIPVTLAGPRSNHNRRPGYKTQNQGPWDEGRVEIDFDYCPGTTVTLSESCGHRGPA
                    TRTTTESGKLITDWCCRSCTLPPLRYQTDSGCWYGMEIRPQRHDEKTLVQSQVNAYNA
                    DMIDPFQLGLLVVFLATQEVLRKRWTAKISMPAILIALLVLVFGGITYTDVLRYVILV
                    GAAFAESNSGGDVVHLALMATFKIQPVFMVASFLKARWTNQENILLMLAAAFFQMAYH
                    DARQILLWEIPDVLNSLAVAWMILRAITFTTTSNVVVPLLALLTPGLRCLNLDVYRIL
                    LLMVGIGSLIREKRSAAAKKKGASLLCLALASTGLFNPMILAAGLIACDPNRKRGWPA
                    TEVMTAVGLMFAIVGGLAELDIDSMAIPMTIAGLMFAAFVISGKSTDMWIERTADISW
                    ESDAEITGSSERVDVRLDDDGNFQLMNDPGAPWKIWMLRMVCLAISAYTPWAILPSVV
                    GFWITLQYTKRGGVLWDTPSPKEYKKGDTTTGVYRIMTRGLLGSYQAGAGVMVEGVFH
                    TLWHTTKGAALMSGEGRLDPYWGSVKEDRLCYGGPWKLQHKWNGQDEVQMIVVEPGKN
                    VKNVQTKPGVFKTPEGEIGAVTLDFPTGTSGSPIVDKNGDVIGLYGNGVIMPNGSYIS
                    AIVQGERMDEPIPAGFEPEMLRKKQITVLDLHPGAGKTRRILPQIIKEAINRRLRTAV
                    LAPTRVVAAEMAEALRGLPIRYQTSAVPREHNGNEIVDVMCHATLTHRLMSPHRVPNY
                    NLFVMDEAHFTDPASIAARGYISTKVELGEAAAIFMTATPPGTSDPFPESNSPISDLQ
                    TEIPDRAWNSGYEWITEYTGKTVWFVPSVKMGNEIALCLQRAGKKVVQLNRKSYETEY
                    PKCKNDDWDFVITTDISEMGANFKASRVIDSRKSVKPTIITEGEGRVILGEPSAVTAA
                    SAAQRRGRIGRNPSQVGDEYCYGGHTNEDDSNFAHWTEARIMLDNINMPNGLIAQFYQ
                    PEREKVYTMDGEYRLRGEERKNFLELLRTADLPVWLAYKVAAAGVPYHDRRWCFDGPR
                    TNTILEDNNEVEVITKLGERKILRPRWIDARVYSDHQALKAFKDFASGKRSQIGLIEV
                    LGKMPEHFMGKTWEALDTMYVVATAEKGGRAHRMALEELPDALQTIALIALLSVMTMG
                    VFFLLMQRKGIGKIGLGGAVLGVATFFCWMAEVPGTKIAGMLLLSLLLMIVLIPEPEK
                    QRSQTDNQLAVFLICVMTLVSAVAANEMGWLDKTKSDISSLFGQRIEVKENFSMGEFL
                    LDLRPATAWSLYAVTTAVLTPLLKHLITSDYINTSLTSINVQASALFTLARGFPFVDV
                    GVSALLLAAGCWGQVTLTVTVTAATLLFCHYAYMVPGWQAEAMRSAQRRTAAGIMKNA
                    VVDGIVATDVPELERATPIMQKKVGQIMLILVSLAAVVVNPSVKTVREAGILITAAAV
                    TLWENGASSVWNATTAIGLCHIMRGGWLSCLSITWTLIKNMEKPGLKRGGAKGRTLGE
                    VWKERLNQMTKEEFTRYRKEAIIEVDRSAAKHARKEGNVTGGHPVSRGTAKLRWLVER
                    RFLEPVGKVIDLGCGRGGWCYYMATQKRVQEVRGYTKGGPGHEEPQLVQSYGWNIVTM
                    KSGVDVFYRPSECCDTLLCDIGESSSSAEVEEHRTIRVLEMVEDWLHRGPREFCVKVL
                    CPYMPKVIEKMELLQRRYGGGLVRNPLSRNSTHEMYWVSRASGNVVHSVNMTSQVLLG
                    RMEKRTWKGPQYEEDVNLGSGTRAVGKPLLNSDTSKIKNRIERLRREYSSTWHHDENH
                    PYRTWNYHGSYDVKPTGSASSLVNGVVRLLSKPWDTITNVTTMAMTDTTPFGQQRVFK
                    EKVDTKAPEPPEGVKYVLNETTNWLWAFLAREKRPRMCSREEFIRKVNSNAALGAMFE
                    EQNQWRSAREAVEDPKFWEMVDEEREAHLRGECHTCIYNMMGKREKKPGEFGKAKGSR
                    AIWFMWLGARFLEFEALGFLNEDHWLGRKNSGGGVEGLGLQKLGYILREVGTRPGGKI
                    YADDTAGWDTRITRADLENEAKVLELLDGEHRRLARAIIELTYRHKVVKVMRPAADGR
                    TVMDVISREDQRGSGQVVTYALNTFTNLAVQLVRMMEGEGVIGPDDVEKLTKGKGPKV
                    RTWLFENGEERLSRMAVSGDDCVVKPLDDRFATSLHFLNAMSKVRKDIQEWKPSTGWY
                    DWQQVPFCSNHFTELIMKDGRTLVVPCRGQDELVGRARISPGAGWNVRDTACLAKSYA
                    QMWLLLYFHRRDLRLMANAICSAVPVNWVPTGRTTWSIHAGGEWMTTEDMLEVWNRVW
                    IEENEWMEDKTPVEKWSDVPYSGKREDIWCGSLIGTRARATWAENIQVAINQVRAIIG
                    DEKYVDYMSSLKRYEDTTLVEDTVL"
    mat_peptide    97..465
                    /product="nucleocapsid protein C"
    mat_peptide    466..741
                    /product="pre-membrane protein prM"
    mat_peptide    742..966
                    /product="membrane protein M"
    mat_peptide    967..2469
                    /product="envelope glycoprotein E"
    mat_peptide    2470..3525
                    /product="non-structural protein 1 NS1"
    mat_peptide    3526..4218
                    /product="non-structural protein 2A NS2A"
    mat_peptide    4219..4611
                    /product="non-structural protein 2B NS2B"
    mat_peptide    4612..6468
                    /product="non-structural protein 3 NS3"
    mat_peptide    6469..6915
                    /product="non-structural protein 4A NS4A"
    mat_peptide    6916..7680
                    /product="non-structural protein NS4B"
    mat_peptide    7681..10395
```

/product="non-structural protein NS5"

```
ORIGIN
        1 agtagttcgc ctgtgtgagc tgacaaactt agtagtgttt gtgaggatta acaacaatta
       61 acacagtgcg agctgtttct tagcacgaag atctcgatgt ctaagaaacc aggagggccc
      121 ggcaagagcc gggctgtcaa tatgctaaaa cgcggaatgc ccgcgtgtt gtccttgatt
      181 ggactgaaga gggctatgtt gagcctgatc gacggcaagg ggccaatacg atttgtgttg
      241 gctctcttgg cgttcttcag gttcacagca attgctccga cccgagcagt gctggatcga
      301 tggagaggtg tgaacaaaca aacagcgatg aaacaccttc tgagtttttaa gaaggaacta
      361 gggaccttga ccagtgctat caatcggcgg agctcaaaac aaaagaaaag aggaggaaag
      421 accggaattg cagtcatgat tggcctgatc gccagcgtag gagcagttac cctctctaac
      481 ttccaaggga aggtgatgat gacggtaaat gctactacg tcacagatgt catcacgatt
      541 ccaacagctg ctggaaagaa cctatgcatt gtcagagcaa tggatgtggg atacatgtgc
      601 gatgatacta tcacttatga atgcccagtg ctgtcggctg gtaatgatcc agaagacatc
      661 gactgttggt gcacaaagtc agcagtctac gtcaggtatg gaagatgcac caagacacgc
      721 cactcaagac gcagtcggag gtcactgaca gtgcagacac acggagaaag cactctagcg
      781 aacaagaagg gggcttggat ggacagcacc aaggccacaa ggtatttggt aaaaacagaa
      841 tcatggatct tgaggaaccc tggatatgcc ctggtggcag ccgtcattgg ttggatgctt
      901 gggagcaaca ccatgcagag agttgtgttt gtcgtgctat tgcttttggt ggccccagct
      961 tacagcttca actgccttgg aatgagcaac agagacttct tggaaggagt gtctggagca
     1021 acatgggtgg atttggttct cgaaggcgac agctgcgtga ctatcatgtc taaggacaag
     1081 cctaccatcg atgtgaagat gatgaatatg gaggcggcca acctggcaga ggtccgcagt
     1141 tattgctatt tggctaccgt cagcgatctc tccaccaaag ctgcgtgccc gaccatggga
     1201 gaagctcaca atgacaaacg tgctgaccca gcttttgtgt gcagacaagg agtggtggac
     1261 aggggctggg gcaacggctg cggactattt ggcaaaggaa gcattgacac atgcgccaaa
     1321 tttgcctgct ctaccaaggc aataggaaga accatcttga aagagaatat caagtacgaa
     1381 gtggccattt ttgtccatgg accaactact gtggagtcgc acggaaacta ctccacacag
     1441 gttggagcca ctcaggcagg agattcagc atcactcctg cggccgcctc atacacacta
     1501 aagcttggag aatatggaga ggtgacagtg gactgtgaac cacggtcagg gattgacacc
     1561 aatgcatact acgtgatgac tgttggaaca aagacgttct tggtccatcg tgagtggttc
     1621 atggacctca acctcccttg gagcagtgct ggaagtactg tgtggaggaa cagagagacg
     1681 ttaatggagt ttgaggaacc acacgccacg aagcagtctg tgatagcatt gggctcacaa
     1741 gagggagctc tgcatcaagc tttggctgga gccattcctg tggaattttc aagcaacact
     1801 gtcaagttga cgtcgggtca tttgaagtgt agagtgaaga tggaaaaatt gcagttgaag
     1861 ggaacaacct atggcgtctg ttcaaaggct ttcaagtttc ttgggactcc cgcagacaca
     1921 ggtcacggca ctgtggtgtt ggaattgcag tacactggca cggatggacc ttgcaaagtt
     1981 cctatctcgt cagtggcttc attgaacgac ctaacgccag tgggcagatt ggtcactgtc
     2041 aaccctttg tttcagtggc cacggccaac gctaaggtcc tgattgaatt ggaaccaccc
     2101 tttggagact catacatagt ggtgggcaga ggagaacaac agatcaatca ccattggcac
     2161 aagtctggaa gcagcattgg caaagccttt acaaccaccc tcaaaggagc gcagagacta
     2221 gccgctctag gagacacagc ttgggacttt ggatcagttg gagggggtgtt cacctcagtt
     2281 gggaaggctg tccatcaagt gttcggagga gcattccgct cactgttcgg aggcatgtcc
     2341 tgg ataacgc aaggattgct gggggctctc ctgttgtgga tgggcatcaa tgctcgtgat
     2401 aggtccatag ctctcacgtt tctcgcagtt ggaggagttc tgctcttcct ctccgtgaac
     2461 gtgcacgctg acactgggtg tgccatagac atcagccggc aagagctgag atgtggaagt
     2521 ggagtgttca tacacaatga tgtggaggct tggatggacc ggtacaagta ttaccctgaa
     2581 acgccacaag gcctagccaa gatcattcag aaagctcata aggaaggagt gtgcggtcta
     2641 cgatcagttt ccagactgga gcatcaaatg tgggaagcag tgaaggacga gctgaacact
     2701 cttttgaagg agaatggtgt ggaccttagt gtcgtggttg agaaacagga gggaatgtac
     2761 aagtcagcac ctaaacgcct caccgccacc acgaaaaat tggaaattgg ctggaaggcc
     2821 tggggaaaga gtatttttatt tgcaccagaa ctcgccaaca cacctttgt ggttgatggt
     2881 ccggagacca aggaatgtcc gactcagaat cgcgcttgga atagcttaga agtggaggat
     2941 tttggatttg gtctcaccag cactcggatg ttcctgaagg tcagagagag caacacaact
     3001 gaatgtgact cgaagatcat tggaacggct gtcaagaaca acttggcgat ccacagtgac
     3061 ctgtcctatt ggattgaaag caggctcaat gatacgtgga agcttgaaag ggcagttctg
     3121 ggtgaagtca aatcatgtac gtggcctgag acgcatacct tgtggggcga tggaatcctt
     3181 gagagtgact tgataatacc agtcacactg gcgggaccac gaagcaatca caatcgggaga
     3241 cctgggtaca gacacaaaa ccaggcccca tgggacgaag gccgggtaga gattgacttc
     3301 gattactgcc caggaactac ggtcaccctg agtgagagct gcggacaccg tggacctgcc
     3361 actcgcacca ccacagagag cggaaagttg ataacagatt ggtgctgcag gagctgcacc
     3421 ttaccaccac tgcgctacca aactgacagc ggctgttggt atggtatgga gatcagacca
     3481 cagagacatg atgaaaagac cctcgtgcag tcacaagtga atgcttataa tgctgatatg
     3541 attgaccctt ttcagttggg ccttctggtc gtgttcttgg ccacccagga ggtccttcgc
     3601 aagaggtgga gcagccaagat cagcatgcca gctatactga ttgctctgct agtcctggtg
     3661 tttggggca ttacttacac tgatgtgtta cgctatgtca tcttggtggg ggcagctttc
     3721 gcagaatcta attcgggagg agacgtggta cacttggcgc tcatggcgac cttcaagata
     3781 caaccagtgt ttatggtggc atcgtttctc aaagcgagat ggaccaacca ggagaacatt
```

```
3841 ttgttgatgt tggcggctgc tttcttttcaa atggcttatc acgatgcccg ccaaattctg
3901 ctctgggaga tccctgatgt gttgaattca ctggcggtag cttggatgat actgagagcc
3961 ataacattca caacgacatc aaacgtggtt gttccgctgc tagccctgct aacacccggg
4021 ctgagatgct tgaatctgga tgtgtacagg atactgctgt tgatggtcgg aataggcagc
4081 ttgatcaggg agaagaggag tgcagctgca aaaagaaag gagcaagtct gctatgcttg
4141 gctctagcct caacaggact tttcaacccc atgatccttg ctgctggact gattgcatgt
4201 gatcccaacc gtaaacgcgg atggcccgca actgaagtga tgacagctgt cggcctaatg
4261 tttgccatcg tcggagggct ggcagagctt gacattgact ccatggccat tccaatgact
4321 atcgcggggc tcatgtttgc tgctttcgtg atttctggga aatcaacaga tatgtggatt
4381 gagagaacgg cggacatttc ctgggaaagt gatgcagaaa ttacaggctc gagcgaaaga
4441 gttgatgtgc ggcttgatga tgatggaaac ttccagctca tgaatgatcc aggagcacct
4501 tggaagatat ggatgctcag aatggtctgt ctcgcgatta gtgcgtacac cccctgggca
4561 atcttgccct cagtagttgg attttggata actctccaat acacaaagag aggaggcgtg
4621 ttgtgggaca ctccctcacc aaaggagtac aaaaaggggg acacgaccac cggcgtctac
4681 aggatcatga ctcgtgggct gctcggcagt tatcaagcag gagcgggcgt gatggttgaa
4741 ggtgtttttcc acacccttg gcatacaaca aaaggagccg ctttgatgag cggagagggc
4801 cgcctggacc catactgggg cagtgtcaag gaggatcgac tttgttacgg aggaccctgg
4861 aaattgcagc acaagtggaa cgggcaggat gaggtgcaga tgattgtggt ggaacctggc
4921 aagaacgtta agaacgtcca gacgaaacca ggggtgttca aaacacctga aggagaaatc
4981 ggggccgtga ctttggactt ccccactgga acatcaggct caccaatagt ggacaaaaac
5041 ggtgatgtga ttgggcttta tggcaatgga gtcataatgc ccaacggctc atacataagc
5101 gcgatagtgc agggtgaaag gatggatgag ccaatcccag ccggattcga acctgagatg
5161 ctgaggaaaa aacagatcac tgtactggat ctccatcccg gcgccggtaa aacaaggagg
5221 attctgccac agatcatcaa agaggccata aacagaagac tgagaacagc cgtgctagca
5281 ccaaccaggg ttgtggctgc tgagatggct gaagcactga gaggactgcc catccggtac
5341 cagacatccg cagtgcccag agaacataat ggaaatgaga ttgttgatgt catgtgtcat
5401 gctaccctca cccacaggct gatgtctcct cacagggtgc cgaactacaa cctgttcgtg
5461 atggatgaag ctcatttcac cgacccagct agcattgcag caagaggtta catttccaca
5521 aaggtcgagc taggggaggc ggcggcaata ttcatgacag ccaccccacc aggcacttca
5581 gatccattcc cagagtccaa ttcaccaatt tccgacttac agactgagat cccggatcga
5641 gcttggaact ctggatacga atggatcaca gaatacaccg gaagacggt ttggtttgtg
5701 cctagtgtca agatggggaa tgagattgcc ctttgcctac aacgtgctgg aaagaaagta
5761 gtccaattga acagaaagtc gtacgagacg gagtacccaa aatgtaagaa cgatgattgg
5821 gactttgtta tcacaacaga catatctgaa atgggggcta acttcaaggc gagcagggtg
5881 attgacagcc ggaagagtgt gaaaccaacc atcataacag aaggagaagg gagagtgatc
5941 ctgggagaac catctgcagt gacagcagct agtgccgccc agagacgtgg acgtatcggt
6001 agaaatccgt cgcaagttgg tgatgagtac tgttatgggg ggcacacgaa tgaagacgac
6061 tcgaacttcg cccattggac tgaggcacga atcatgctgg acaacatcaa catgccaaac
6121 ggactgatcg ctcaattcta ccaaccagag cgtgagaagg tatataccat ggatgggggaa
6181 taccggctca gaggagaaga gagaaaaaac tttctggaac tgttgaggac tgcagatctg
6241 ccagtttggc tggcttacaa ggttgcagcg gctggagtgc cataccacga ccggaggtgg
6301 tgctttgatg gtcctaggac aaacacaatt ttagaagaca acaacgaagt ggaagtcatc
6361 acgaagcttg gtgaaaggaa gattctgagg ccgcgctgga ttgacgccag ggtgtactcg
6421 gatcaccagg cactaaaggc gttcaaggac ttcgcctcgg gaaaacgttc tcagataggg
6481 ctcattgagg ttctgggaaa gatgcctgag cacttcatgg ggaagacatg ggaagcactt
6541 gacaccatgt acgttgtggc cactgcagag aaaggaggaa gagctcacag aatggccctg
6601 gaggaactgc cagatgctct tcagacaatt gccttgattg ccttattgag tgtgatgacc
6661 atgggagtat tcttcctcct catgcagcgg aagggcattg gaaagatagg tttgggaggc
6721 gctgtcttgg gagtcgcgac ctttttctgt tggatggctg aagttccagg aacgaagatc
6781 gccggaatgt tgctgctctc ccttctcttg atgattgtgc taattcctga gccagagaag
6841 caacgttcgc agacagacaa ccagctagcc gtgttcctga tttgtgtcat gacccttgtg
6901 agcgcagtgg cagccaacga gatggggttgg ctagataaga ccaagagtga cataagcagt
6961 ttgtttgggc aaagaattga ggtcaaggag aatttcagca tgggagagtt tcttctggac
7021 ttgaggccgg caacagcctg gtcactgtac gctgtgacaa cagcggtcct cactccactg
7081 ctaaagcatt tgatcacgtc agattacatc aacacctcat tgacctcaat aaacgttcag
7141 gcaagtgcac tattcacact cgcgcgaggc ttcccccttcg tcgatgttgg agtgtcggct
7201 ctcctgctag cagccggatg ctgggggacaa gtcacccctca ccgttacggt aacagcggca
7261 acactccttt tttgccacta tgcctacatg gttcccggtt ggcaagctga ggcaatgcgc
7321 tcagcccagc ggcggacagc ggccggaatc atgaagaacg ctgtagtgga tggcatcgtg
7381 gccacggacg tcccagaatt agagcgcgcc acacccatca tgcagaagaa agttggacag
7441 atcatgctga tcttggtgtc tctagctgca gtagtagtga acccgtctgt gaagacagta
7501 cgagaagccg gaattttgat cacggccgca gcggtgacgc tttgggagaa tggagcaagc
7561 tctgtttgga acgcaacaac tgccatcgga ctctgccaca tcatgcgtgg gggttggttg
7621 tcatgtctat ccataacatg gacactcata aagaacatgg aaaaaccagg actaaaaaga
7681 ggtggggcaa aaggacgcac cttgggagag gtttggaaag aaagactcaa ccagatgaca
7741 aaagaagagt tcactaggta ccgcaaagag gccatcatcg aagtcgatcg ctcagcggca
```

```
 7801 aaacacgcca ggaaagaagg caatgtcact ggagggcatc cagtctctag gggcacagca
 7861 aaactgagat ggctggtcga acggaggttt ctcgaaccgg tcggaaaagt gattgacctt
 7921 ggatgtggaa gaggcggttg gtgttactat atggcaaccc aaaaaagagt ccaagaagtc
 7981 agagggtaca caaagggcgg tcccggacat gaagagcccc aactagtgca aagttatgga
 8041 tggaacattg tcaccatgaa gagtggagtg gatgtgttct acagaccttc tgagtgttgt
 8101 gacaccctcc tttgtgacat cggagagtcc tcgtcaagtg ctgaggttga agagcatagg
 8161 acgattcggg tccttgaaat ggttgaggac tggctgcacc gagggccaag ggaattttgc
 8221 gtgaaggtgc tctgccccta catgccgaaa gtcatagaga agatggagct gctccaacgc
 8281 cggtatgggg ggggactggt cagaaaccca ctctcacgga attccacgca cgagatgtat
 8341 tgggtgagtc gagcttcagg caatgtggta cattcagtga atatgaccag ccaggtgctc
 8401 ctaggaagaa tggaaaaaag gacctggaag ggaccccaat acgaggaaga tgtaaacttg
 8461 ggaagtggaa ccagggcggt gggaaaaccc ctgctcaact cagacaccag taaaatcaag
 8521 aacaggattg aacgactcag gcgtgagtac agttcgacgt ggcaccacga tgagaaccac
 8581 ccatatagaa cctggaacta tcacggcagt tatgatgtga agcccacagg ctccgccagt
 8641 tcgctggtca atggagtggt caggctcctc tcaaaaccat gggacaccat cacgaatgtt
 8701 accaccatgg ccatgactga cactactccc ttcgggcagc agcgagtgtt caaagagaag
 8761 gtggacacga aagctcctga accgccagaa ggagtgaagt acgtgctcaa cgagaccacc
 8821 aactggttgt gggcgttttt ggccagagaa aaacgtccca gaatgtgctc tcgagaggaa
 8881 ttcataagaa aggtcaacag caatgcagct ttgggtgcca tgtttgaaga gcagaatcaa
 8941 tggaggagcg ccagagaagc agttgaagat ccaaaatttt gggagatggt ggatgaggag
 9001 cgcgaggcac atctgcgggg ggaatgtcac acttgcattt acaacatgat gggaaagaga
 9061 gagaaaaaac ccggagagtt cggaaaggcc aagggaagca gagccatttg gttcatgtgg
 9121 ctcggagctc gctttctgga gttcgaggct ctgggttttc tcaatgaaga ccactggctt
 9181 ggaagaaaga actcaggagg aggtgtcgag ggcttgggcc tccaaaaact gggttacatc
 9241 ctgcgtgaag ttggcacccg gcctggggg aagatctatg ctgatgacac agctggctgg
 9301 gacacccgca tcacgagagc tgacttggaa aatgaagcta aggtgcttga gctgcttgat
 9361 ggggaacatc ggcgtcttgc cagggccatc attgagctca cctatcgtca caagttgtg
 9421 aaagtgatgc gccgggctgac tgatggaaga accgtcatgg atgttatctc cagagaagat
 9481 cagagggggga gtggacaagt tgtcacctac gccctaaaca ctttcaccaa cctggccgtc
 9541 cagctggtga ggatgatgga aggggaagga gtgattggcc cagatgatgt ggagaaactc
 9601 acaaaaggga aaggacccaa agtcaggacc tggctgtttg agaatgggga agaaagactc
 9661 agccgcatgg ctgtcagtgg agatgactgt gtggtaaagc ccctggacga tcgctttgcc
 9721 acctcgctcc acttcctcaa tgctatgtca aaggttcgca aagacatcca agagtggaaa
 9781 ccgtcaactg gatggtatga ttggcagcag gttccatttt gctcaaacca tttcactgaa
 9841 ttgatcatga aagatggaag aacactggtg gttccatgcc gaggacagga tgaattggta
 9901 ggcagagctc gcatatctcc aggggccgga tggaacgtcc gcgacactgc ttgtctggct
 9961 aagtcttatg cccagatgtg gctgcttctg tacttccaca gaagagacct gcggctcatg
10021 gccaacgcca tttgctccgc tgtccctgtg aattgggtcc ctaccggaag aaccacgtgg
10081 tccatccatg caggaggaga gtggatgaca acagaggaca tgttggaggt ctggaaccgt
10141 gtttggatag aggagaatga atggatggaa gacaaaaccc cagtggagaa atggagtgac
10201 gtcccatatt caggaaaacg agaggacatc tggtgtggca gcctgattgg cacaagagcc
10261 cgagccacgt gggcagaaaa catccaggtg gctatcaacc aagtcagagc aatcatcgga
10321 gatgagaagt atgtggatta catgagttca ctaaagagat atgaagacac aactttggtt
10381 gaggacacag tactgtagat atttaatcaa ttgtaaatag acaatataag tatgcataaa
10441 agtgtagttt tatagtagta tttagtggtg ttagtgtaaa tagttaagaa aattttgagg
10501 agaaagtcag gccgggaagt tcccgccacc ggaagttgag tagacggtgc tgcctgcgac
10561 tcaaccccag gaggactggg tgaacaaagc cgcgaagtga tccatgtaag ccctcagaac
10621 cgtctcggaa ggaggacccc acatgttgta acttcaaagc ccaatgtcag accacgctac
10681 ggcgtgctac tctgcggaga gtgcagtctg cgatagtgcc ccaggaggac tgggttaaca
10741 aaggcaaacc aacgccccac gcggccctag ccccggtaat ggtgttaacc agggcgaaag
10801 gactagaggt tagaggagac cccgcggttt aaagtgcacg gcccagcctg gctgaagctg
10861 taggtcaggg gaaggactag aggttagtgg agaccccgtg ccacaaaaca ccacaacaaa
10921 acagcatatt gacacctggg atagactagg agatcttctg ctctgcacaa ccagccacac
10981 ggcacagtgc gccgacaatg gtggctggtg gtgcgagaac acaggatct
//
```

Apr 17 2007 11:10:07

| PubMed | Nucleotide | Protein | Genome | Structure | PMC | Taxonomy | OMIM | Books |

Search CoreNucleotide for [   ] Go Clear

Limits   Preview/Index   History   Clipboard   Details

Display GenBank   Show 5   Send to   Hide: ☐ sequence   ☐ all but gene, CDS and mRNA features

Range: from begin   to end   ☐ Reverse complemented strand   Features: [+] Refresh

☐ **1: AF260968.** Reports West Nile virus s...[gi:9930135]   Links

Features   Sequence

```
LOCUS       AF260968              11029 bp    RNA     linear   VRL 27-AUG-2000
DEFINITION  West Nile virus strain Eg101, complete genome.
ACCESSION   AF260968
VERSION     AF260968.1  GI:9930135
KEYWORDS    .
SOURCE      West Nile virus
  ORGANISM  West Nile virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
            Flavivirus; Japanese encephalitis virus group.
REFERENCE   1  (bases 1 to 11029)
  AUTHORS   Bowen,M., Meyer,R.F., McKinney,N., Morrill,W. and Lanciotti,R.
  TITLE     Complete genomic sequence of West Nile virus strain Eg101
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 11029)
  AUTHORS   Bowen,M., Meyer,R.F., McKinney,N., Morrill,W. and Lanciotti,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-APR-2000) Arbovirus Diseases Branch, Centers for
            Disease Control & Prevention, Rampart Road, Fort Collins, CO 80521,
            USA
FEATURES             Location/Qualifiers
     source          1..11029
                     /organism="West Nile virus"
                     /mol_type="genomic RNA"
                     /strain="Eg101"
                     /db_xref="taxon:11082"
     CDS             97..10398
                     /codon_start=1
                     /product="polyprotein precursor"
                     /protein_id="AAG02039.1"
                     /db_xref="GI:9930136"
                     /translation="MSKKPGGPGKSRAVNMLKRGMPRVLSLIGLKRAMLSLIDGKGPI
                     RFVLALLAFFRFTAIAPTRAVLDRWRGVNKQTAMKHLLSFKKELGTLTSAINRRSSKQ
                     KKRGGKTGIAVMIGLIASVGAVTLSNFQGKVMMTVNATDVTDVITIPTAAGKNLCIVR
                     AMDVGYMCDDTITYECPVLSAGNDPEDIDCWCTKSAVYVRYGRCTKTRHSRRSRRSLT
                     VQTHGESTLANKKGAWMDSTKATRYLVKTESWILRNPGYALVAAVIGWMLGSNTMQRV
                     VFVVLLLLVAPAYSFNCLGMSNRDFLEGVSGATWVDLVLEGDSCVTIMSKDKPTIDVK
                     MMNMEAANLAEVRSYCYLATVSDLSTKAACPTMGEAHNDKRADPAFVCKQGVVDRGWG
                     NGCGLFGKGSIDTCAKFACSTKATGRTILKENIKYEVAIFVHGPTTVESHGNYPTQIG
                     ATQAGRFSITPAAPSYTLKLGEYGEVTVDCEPRSGIDTNAYVMTVGTKTFLVHREWF
                     MDLNLPWSSAGSTVWRNRETLMEFEEPHATKQSVIALGSQEGALHQALAGAIPVEFSS
                     NTVKLTSGHLKCRVKMEKLQLKGTTYGVCSKAFKFLGTPADTGHGTVVLELQYTGTDG
                     PCKVPISSVASLNDLTPVGRLVTVNPFVSVATANAKVLIELEPPFGDSYIVVGRGEQQ
                     INHHWHKSGSSIGKAFTTTLKGAQRLAALGDTAWDFGSVGGVFTSVGKAVHQVFGGAF
                     RSLFGGMSWITQGLLGALLLWMGINARDRSIALTFLAVGGVLLFLSVNVHADTGCAID
                     ISRQELRCGSGVFIHNDVEAWMDRYKYYPETPQGLAKIIQKAHKEGVCGLRSVSRLEH
                     QMWEAVKDELNTLLKENGVDLSVVVEKQEGMYKSAPKRLTATTEKLEIGWKAWGKSIL
                     FAPELANNTFVVDGPETKECPTQNRAWNSLEVEDFGFGLTSTRMFLKVRESNTTECDS
                     KIIGTAVKNNLAIHSDLSYWIESRLNDTWKLERAVLGEVKSCTWPETHTLWGEGILES
                     DLIIPVTLAGPRSNHNRRPGYKTQNQGPWDEGRVEIDFDYCPGTTVTLSESCGHRGPA
```

```
                    TRTTTESGKLITDWCCRSCTLPPLRYQTDSGCWYGMEIRPQRHDEKTLVQSQVNAYNA
                    DMIDPFQLGLLVVFLATQEVLRKRWTAKISMPAILIALLVLVFGGITYTDVLRYVILV
                    GAAFAESNSGGDVVHLALMATFKIQPVFMVASFLKARWTNQENILLMLAAVFFQMAYH
                    DARQILLWEIPDVLNSLAVAWMILRAITFTTTSNVVVPLLALLTPGLRCLNLDVYRIL
                    LLMVGIGSLIREKRSAAAKKKGASLLCLALASTGLFNPMILAAGLIACDPNRKRGWPA
                    TEVMTAVGLMFAIVGGLAELDIDSMAIPMTIAGLMFAAFVISGKSTDMWIERTADISW
                    ESDAEITGSSERVDVRLDDDGNFQLMNDPGAPWKIWMLRMACLAISAYTPWAILPSVV
                    GFWITLQYTKRGGVLWDTPSPKEYKKGDTTTGVYRIMTRGLLGSYQAGAGVMVEGVFH
                    TLWHTTKGAALMSGEGRLDPYWGSVKEDRLCYGGPWKLQHKWNGQDEVQMIVVEPGKN
                    VKNVQTKPGVFKTPEGEIGAVTLDFPTGTSGSPIVDKNGDVIGLYGNGVIMPNGSYIS
                    AIVQGERMDEPIPAGFEPEMLRKKQITVLDLHPGAGKTRRILPQIIKEAINRRLRTAV
                    LAPTRVVAAEMAEALRGLPIRYQTSAVPREHNGNEIVDVMCHATLTHRLMSPHRVPNY
                    NLFVMDEAHFTDPASIAARGYISTKVELGEAAAIFMTATPPGTSDPFPESNSPISDLQ
                    TEIPDRAWNSGYEWITEYIGKTVWFVPSVKMGNEIALCLQRAGKKVVQLNRKSYETEY
                    PKCKNDDWDFVITTDISEMGANFKASRVIDSRKSVKPTIITEGEGRVILGEPSAVTAA
                    SAAQRRGRIGRNPSQVGDEYCYGGHTNEDDSNFAHWTEARIMLDNINMPNGLIAQFYQ
                    PEREKVYTMDGEYRLRGEERKNFLELLRTADLPVWLAYKVAAAGVSYHDRRWCFDGPR
                    TNTILEDNNEVEVITKLGERKILRPRWIDARVYSDHQALKAFKDFASGKRSQIGLIEV
                    LGKMPEHFMGKTWEALDTMYVVATAEKGGRAHRMALEEELPDALQTIALIALLSVMTMG
                    VFFLLMQRKGIGKIGLGGVVLGVATFFCWMAEVPGTKIAGMLLLSLLLMIVLIPEPEK
                    QRSQTDNQLAVFLICVLTLVSAVAANEMGWLDKTKNDISSLFGQRIEAKENFSMGEFL
                    LDLRPATAWSLYAVTTAVLTPLLKHLITSDYINTSLTSINVQASALFTLARGFPFVDV
                    GVSALLLAAGCWGQVTLTVTVTAATLLFCHYAYMVPGWQAEAMRSAQRRTAAGIMKNA
                    VVDGIVATDVPELERTTPIMQKKVGQIMLILVSLAAVVVNPSVKTVREAGILITAAAV
                    TLWENGASSVWNATTAIGLCHIMRGGWLSCLSITWTLIKNMEKPGLKRGGAKGRTLGE
                    VWKERLNQMTKEEFTRYRKEAIIEVDRSAAKHARKEGNVTGGHPVSRGTAKLRWLVER
                    RFLEPVGKVIDLGCGRGGWCYYMATQKRVQEVRGYTKGGPGHEEPQLVQSYGWNIVTM
                    KSGVDVFYRPSECCDTLLCDIGESSSSAEVEEHRTIRVLEMVEDWLHRGPKEFCVKVL
                    CPYMPKVIEKMELLQRRYGGGLVRNPLSRNSTHEMYWVSRASGNVVHSVNMTSQVLLG
                    RMEKRTWKGPQYEEDVNLGSGTRAVGKPLLNSDTSKIKNRIERLRREYSSTWHHDENH
                    PYRTWNYHGSYDVKPTGSASSLVNGVVRLLSKPWDTITNVTTMAMTDTTPFGQQRVFK
                    EKVDTKAPEPPEGVKYVLNETTNWLWAFLAREKRPRMCSREEFIKKVNSNAALGAMFE
                    EQNQWRSAREAVEDPKFWEMVDEEREAHLRGECHTCIYNMMGKREKKPGEFGKAKGSR
                    AIWFMWLGARFLEFEALGFLNEDHWLGRKNSGGGVEGLGLQKLGYILREVGTRPGGKI
                    YADDTAGWDTRITRADLENEAKVLELLDGEHRRLARAIIELTYRHKVVKVMRPAADGR
                    TVMDVISREDQRGSGQVVTYALNTFTNLAVQLVRMMEGEGVIGPDDVEKLTKGKGPKV
                    RTWLFENGEERLSRMAVSGDDCVVKPLDDRFATSLHFLNAMSKVRKDIQEWKPSTGWY
                    DWQQVPFCSNHFTELIMKDGRTLVVPCRGQDELVGRARISPGAGWNVRDTACLAKSYA
                    QMWLLLYFHRRDLRLMANAICSAVPVNWVPTGRTTWSIHAGGEWMTTEDMLEVWNRVW
                    IEENEWMEDKTPVEKWSDVPYSGKREDIWCGSLIGTRTRATWAENIQVAINQVRAIIG
                    DEKYVDYMSSLKRYEDTTLVEDTVL"
    mat_peptide    97..465
                    /product="nucleocapsid protein C"
    mat_peptide    466..741
                    /product="pre-membrane protein prM"
    mat_peptide    742..966
                    /product="membrane protein M"
    mat_peptide    967..2469
                    /product="envelope glycoprotein E"
    mat_peptide    2470..3525
                    /product="non-structural protein 1 NS1"
    mat_peptide    3526..4218
                    /product="non-structural protein 2A NS2A"
    mat_peptide    4219..4611
                    /product="non-structural protein 2B NS2B"
    mat_peptide    4612..6468
                    /product="non-structural protein 3 NS3"
    mat_peptide    6469..6915
                    /product="non-structural protein 4A NS4A"
    mat_peptide    6916..7680
                    /product="non-structural protein NS4B"
    mat_peptide    7681..10395
                    /product="non-structural protein NS5"
ORIGIN
        1 agtagttcgc ctgtgtgagc tgacaaactt agtagtgttt gtgaggatta acaacaatta
       61 acacggtgcg agctgtttct tagcacgaag atctcgatgt ctaagaaacc aggagggccc
```

```
 121 ggcaagagcc gggctgtcaa tatgctaaaa cgcggaatgc cccgcgtgtt gtccttgatt
 181 ggactgaaga gggcaatgtt gagcctgatc gacggcaagg gaccaatacg atttgtgttg
 241 gctctcttgg cgttcttcag gttcacagca attgctccga cccgagcagt gctggatcga
 301 tggagaggtg tgaacaaaca aacagcgatg aaacaccttc tgagtttaa gaaggaacta
 361 gggaccttga ccagtgctat caatcggcgg agctcaaaac aaaagaaaag aggaggaaag
 421 accggaattg cagtcatgat tggcttgatc gccagcgtgg gagcagttac cctctctaac
 481 ttccaaggga aggtgatgat gactgtaaat gccactgacg tcacagacgt catcacgatt
 541 ccaacagctg ctggaaagaa tctatgcatt gtcagagcaa tggacgtggg gtacatgtgt
 601 gatgtatacta tcacctatga atgtccagtg ctgtcggctg gtaatgatcc agaagacatc
 661 gactgttggt gcacaaaatc agcagtctac gtcaggtatg gaagatgcac caagacacgc
 721 cactcaagac gtagccggag gtcactgaca gtgcagacac atggagaaag cactctagcg
 781 aacaagaagg gggcttggat ggacagcacc aaggctacaa ggtatttggt aaaaacagaa
 841 tcatggatct tgaggaaccc cggatatgcc ctggtggcag ccgtcattgg ttggatgctt
 901 ggaagcaaca ccatgcagcg agttgtgttc gttgtgctac tgctcttggt ggctccagcc
 961 tacagcttta actgccttgg aatgagcaac agagacttct tagagggagt gtctggagca
1021 acatgggtgg atttggttct cgaaggcgac agctgtgtga ccatcatgtc taaggacaag
1081 cctaccatcg atgtgaagat gatgaatatg gaggccgcca acctggcaga ggtccgcagt
1141 tattgctatc tggccaccgt cagcgatctc tccaccaaag ctgccgtgccc gactatggga
1201 gaagctcaca atgacaaacg tgctgaccca gcttttgtgt gtaaacaagg agtagtggac
1261 aggggttggg gcaacggctg tggactattt ggtaaaggaa gcattgacac atgcgccaaa
1321 tttgcctgtt ctaccaaggc aacaggaaga accattctga agagaacat caagtacgaa
1381 gtggctatct ttgtccatgg accaaccact gtggagtcgc atggaaacta ccccacacag
1441 attgggggcca ctcaggcagg gagattcagc atcactcctg cggcgccttc atacacacta
1501 aaacttggag agtatggaga ggtgacggtg gactgtgaac cacgatcagg gattgacacc
1561 aatgcatact acgtgatgac tgtcggaaca aagacgttct tggtccatcg tgagtggttt
1621 atggacctca acctcccctg gagcagtgcc ggaagcactg tgtggaggaa cagagagacg
1681 ttgatggagt ttgaggaacc acacgccacg aagcagttct gatagcatt gggctcacaa
1741 gagggagctc tgcatcaagc tttggctgga gccattcctg tggaatttc aagcaacact
1801 gtcaagttga catcgggtca tttgaagtgt agagtgaaga tggaaaaatt gcagttgaag
1861 ggaacaacct acggcgtctg ttcaaaggct ttcaagtttc ttggaactcc cgcagacaca
1921 ggccacggca ctgtagtgtt ggaattgcag tacactggca cggatggacc ttgcaaagtt
1981 cccatctcgt cagtggcttc attgaacgac ctaacgccag tgggcaggtt ggtcactgtc
2041 aacccctttg tttcagtagc cacggccaat gccaaggtcc tgattgaatt ggaaccaccc
2101 tttggagact catacatagt ggtgggcaga ggagaacaac agattaatca ccattggcac
2161 aagtctggaa gcagcattgg caaagccttc acaaccaccc tcaaaggggc gcagagatta
2221 gccgccctag gagatacagc ttgggacttt ggatcagttg gagggggtgtt cacctcagtg
2281 gggaaggctg tccatcaagt gtttggtgga gcattccgct cactgttcgg aggcatgtct
2341 tggataacgc aaggattgct gggggctctg ctgttgtgga tgggcatcaa tgctcgtgac
2401 aggtccatag ctctcacgtt tctcgcagtt ggaggggttt tgctctttct ctccgtgaac
2461 gtgcacgctg acactggatg tgccatagac atcagccggc aggagctgag atgtggaagt
2521 ggagtgttca tacacaatga tgtggaggct tggatggacc ggtacaagta ctaccctgaa
2581 acgccacaag gcctagccaa gatcattcaa aaagcccaca aagaaggagt gtgcggtcta
2641 cggtcggttt ccagactgga gcaccaaatg tgggaagcgg tgaaggacga gctaaacact
2701 cttttgaaag agaatggtgt ggacctcagt gttgtggttg agaaacagga gggaatgtac
2761 aagtcagcac ctaaacgtct caccgctacc acggaaaaat tggaaatagg ctggaaggcc
2821 tggggaaaga gcatcctatt cgcaccagaa ttggccaaca cacttttgt ggttgatggt
2881 ccggagacca aggaatgccc aactcagaat cgcgcttgga acagcttgga agtagaggat
2941 tttggatttg gtctcaccag tacccggatg ttcctgaagg tcagagagag caacacaact
3001 gaatgtgact caaagatcat cggaacggct gtcaagaaca acttggcgat ccacagtgac
3061 ctgtcctatt ggattgaaag caggcttaat gatacgtgga agcttgaaag ggcggtcctg
3121 ggtgaagtta aatcatgcac ttggcctgaa acgcacactt tgtggggtga aggaatcctc
3181 gagagtgact tgataatacc agtcacactg gcgggaccac gaagcaacca caatcggaga
3241 cctgggtaca agacacaaaa ccagggccca tgggacgaag gccgggtaga gattgattc
3301 gattactgcc caggaacgac ggtcaccctg agtgagagct gcggacaccg tggacctgcc
3361 actcgcacca ccacagagag cggaaagctg ataacggact ggtgctgcag gagctgcacc
3421 ttaccaccat gcgctacca gacggacagc ggttgttggt atggtatgga gattagacca
3481 cagaggcatg atgaaaagac ccttgtgcag tcacaagtga atgcttacaa cgctgatatg
3541 attgatcctt ttcagctggg ccttctggtc gtgttcttgg ccacccagga ggtccttcgc
3601 aagaggtgga cagccaagat cagcatgcca gctatactga ttgctctgct agtcctggtg
3661 tttgggggca ttacttacac tgacgtgtta cgctatgtca tcttagtggg agcagctttc
3721 gcagaatcca attcgggagg agacgtggta cacttggcgc tcatggcgac cttcaagata
3781 caaccagtgt ttatggtggc atcgtttctc aaagcgagat ggaccaacca ggagaatatc
3841 ttgttgatgt tggcggctgt tttctttcaa atggcttacc atgacgctcg ccaaattctg
3901 ctttgggaga tccctgatgt gttgaattca ttggcagtag cttggatgat actgagagcc
3961 ataacctta caacaacatc aaacgtggtt gttccgctgc tagctctgtt aacacccgga
4021 ctgagatgct tgaatctgga tgtgtacagg atcctgctat tgatggtcgg aataggcagc
```

```
4081 ttgatcagag agaagagaag cgcagctgca aaaaagaaag gagcaagtct gttatgcctg
4141 gctctagcct caacaggact tttcaaccct atgatcctcg ccgctggact cattgcatgt
4201 gatcccaacc gtaaacgagg atggcccgca actgaagtga tgactgctgt cggcctgatg
4261 tttgccattg tcggagggct ggcagagctt gacattgact ccatggccat tccaatgacc
4321 atcgcagggc tcatgtttgc tgccttcgtg atatctggga aatcaacaga tatgtggatc
4381 gagaggacgg cggacatctc ctgggaaagt gatgcggaaa ttacaggctc gagcgagaga
4441 gttgatgtgc ggcttgatga tgacggaaat ttccagctca tgaatgatcc aggagcacct
4501 tggaagatat ggatgctcag aatggcttgc ctcgcgatta gtgcgtacac cccttgggca
4561 atcctgccct cagtagttgg attttggata actctccaat acacaaagag aggaggtgtg
4621 ctgtgggaca ctccctcacc aaaggagtac aaaaaagggg acacgaccac tggcgtctac
4681 aggatcatga ctcgtgggct gctcggcagt tatcaagcag gagcgggcgt gatggttgaa
4741 ggggtttttcc acacccttttg gcatacaaca aaaggagccg ctctgatgag cggggaaggc
4801 cgcctggacc catactgggg tagtgtcaaa gaggatcgac tttgctacgg aggaccctgg
4861 aaattgcagc acaagtggaa tgggcaggat gaggtgcaaa tgattgtggt ggaacctggc
4921 aagaacgtta aaaacgtcca gacgaaacca ggggtgttca aaacacctga aggagaaatt
4981 ggggccgtga ctctggactt ccccactgga acatcaggct caccaatagt ggacaaaaac
5041 ggtgatgtga tcgggctcta tggcaatgga gtcataatgc ccaacggctc atacataagc
5101 gcgatagtgc agggtgaaag gatggatgag ccgatcccag ccggattcga acctgagatg
5161 ctgaggaaaa aacagatcac agttctggac cttcatcccg gtgctggtaa aacaaggagg
5221 atactgccac agatcatcaa agaggccata aatagaagat tgagaacggc cgtgctagca
5281 ccaactaggg ttgtagccgc tgagatggct gaagccctga gaggactgcc catccggtat
5341 cagacatctg cagtgcccag agaacacaat ggaaatgaga ttgttgatgt catgtgccat
5401 gccactctca ctcacaggct gatgtctcct cacagggtgc cgaactacaa tcttttcgtg
5461 atggatgagg ctcattttac cgacccagct agcattgcag caaggggtta tatttccaca
5521 aaagtcgagc tggggggaggc ggcggcaata ttcatgacag ctaccccacc aggcacttca
5581 gacccattcc cagagtccaa ttcacctatt tctgacttgc agactgagat cccagatcgg
5641 gcctggaact ctgggtacga atggattaca gaatacattg ggaaaacggt ttggtttgtg
5701 cccagtgtga aaatggggaa tgagattgcc ctttgtctac aacgtgccgg caaaaaagta
5761 gtccaactga acagaaagtc gtatgagacg gagtacccaa agtgcaagaa cgatgattgg
5821 gactttgtta tcacaacaga catatctgaa atgggggcta acttcaaggc gagcagggtg
5881 attgacagca ggaagagtgt gaaaccaacc atcatcacgg aaggagaagg gagggtgatc
5941 ctgggagaac catccgctgt gacagcagct agtgcagccc aaagacgtgg acgcatcggt
6001 aggaatccat cgcaagttgg tgatgagtac tgctatgggg ggcacacgaa tgaagacgac
6061 tcgaacttcg cccattggac tgaggcacga atcatgctgg acaacatcaa catgccaaac
6121 ggactgatcg ctcaattcta ccaaccgagg cgtgaaaagg tatacaccat ggatggagaa
6181 taccgactca gaggagaaga gaggaaaaac tttctggaat tattgaggac tgcagatctg
6241 ccagtttggc tggcttacaa ggtggcagcg gctggagtgt cataccacga tcggagatgg
6301 tgttttgatg ccctaggac aaacacaatt ctagaagaca caacgaagt ggaagtcatt
6361 acgaagcttg gtgaaagaaa gattctgagg ccgcgctgga ttgacgccag ggtgtactcg
6421 gatcatcagg cattaaaggc gttcaaggac tttgcttcgg gaaagcgttc tcagataggg
6481 ctcattgagg ttctgggaaa gatgcctgag cacttcatgg ggaagacatg ggaagcactt
6541 gacaccatgt atgttgtggc caccgcagag aaaggggggaa gagctcacag aatggccttg
6601 gaggaactgc cagatgctct ccagacaatt gccctgattg ccttattgag tgtgatgacc
6661 atgggagtat tcttcctcct catgcagcgg aagggcattg gaaagatagg tttggggaggc
6721 gttgtcctgg gagtcgcaac cttctttttgt tggatggctg aagttccagg aacgaagatc
6781 gccggaatgt tgctgctttc ccttctcttg atgattgtgc taatccctga gccagagaag
6841 caacgttcgc agacagacaa ccagctagcc gtgttcctga tttgtgtgtt gaccctcgtg
6901 agcgcagtgg cagccaacga aatgggttgg ctggacaaga ccaagaatga tataagcagt
6961 ttgtttgggc aaagaattga ggccaaggag aatttcagta tgggagagtt tctcctggac
7021 ttgagaccgg caacagcctg gtcactgtat gctgtgacca cagcggttct cactccactg
7081 ctaaagcatc tgatcacgtc agattacatc aacacttcat tgacctcaat caatgttcaa
7141 gcaagtgcac tattcacact cgcgcgaggc ttcccctttg tcgatgttgg agtgtcggct
7201 ctcctgctag cagccggatg ctgggacaa gtcaccctca ccgtgacggt gacagcggca
7261 acactcctgt tctgccacta cgcctacatg gttcccggat ggcaggctga ggcaatgcgc
7321 tcagcccagc ggcggacagc ggctggaatc atgaaaaacg ctgtagtgga tggcatcgtg
7381 gccacggacg tcccagaatt agagcgcacc acacccatca tgcagaagaa agttgggcaa
7441 atcatgctcga tcttggtgtc tctagctgca gtagtgctga acccgtctgt gaagacagtg
7501 cgagaagccg gaattctgat cacggcagca gcggtgacac tctggagaga tggagcaagc
7561 tctgtttgga atgcaacaac tgccatcgga ctctgccaca tcatgcgtgg gggttggttg
7621 tcatgcttat ccataacatg gacactcata agaacatgg aaaaaccagg actaaaaaga
7681 ggtggggcaa agggacgcac cttgggagag gttttggaaag aaagactcaa ccagatgaca
7741 aaagaagagt tcactaggta ccgcaaagag gccatcatcg aagtcgatcg ctcagcagca
7801 aaacacgcca ggaaagaagg caatgtcact ggagggcatc cagtctctag aggcacagca
7861 aagctgagat ggctggtcga gcggaggttt ctcgaaccgg tcggaaaagt gattgacctt
7921 ggatgtggaa gaggcggttg gtgttactac atggcaaccc aaaaaagagt ccaagaggtc
7981 agagggtaca caaagggtgg tcccggacat gaagagcccc aactggtgca aagttatgga
```

```
 8041 tggaacattg tcaccatgaa gagcggagtg gatgtgttct acagaccttc tgagtgctgc
 8101 gataccctcc tttgtgacat cggagagtct tcatcaagtg ctgaggttga agagcatagg
 8161 acgatccggg tccttgaaat ggttgaggac tggctgcacc gagggccaaa ggaattttgt
 8221 gtgaaggtgc tctgccccta tatgccaaaa gtcatagaaa agatggagct gctccagcgc
 8281 cggtatgggg ggggactggt cagaaaccca ctctcgcgga attccacgca cgagatgtat
 8341 tgggtaagtc gagcttcggg caatgtggta cactcagtga acatgaccag ccaggtgctt
 8401 ctgggaagaa tggagaaaag gacctggaag ggaccccaat acgaggaaga tgtgaacttg
 8461 ggaagtggaa ccagggcggt gggaaaaccc ctactcaact cagacactag taaaatcaag
 8521 aacaggattg aacgactcag gcgtgagtac agttcgacgt ggcaccacga tgagaaccac
 8581 ccatatagaa cctggaacta tcacggcagt tatgatgtga aacctacagg ctccgccagc
 8641 tcgctggtca atggagtggt taggctcctc tcaaaaccat gggacaccat cacgaacgtt
 8701 accaccatgg ccatgactga cactactccc ttcggacagc agcgggtgtt taaagagaag
 8761 gtggacacga aagctcctga accgccagaa ggagtgaagt atgtgctcaa tgaaaccacc
 8821 aactggttgt gggcgtttct ggccagagaa aaacgtccca gaatgtgctc tcgagaggaa
 8881 ttcataaaaa aggtcaatag caatgcagct ctgggtgcca tgtttgaaga gcagaaccaa
 8941 tggaggagcg ccagagaagc agttgaggat ccaaaatttt gggagatggt ggatgaggag
 9001 cgcgaggcac acctgcgggg ggaatgtcac acttgcatct acaacatgat ggggaagaga
 9061 gagaagaaac ctggagagtt cggaaaggct aagggaagca gagccatatg gttcatgtgg
 9121 ctcggagctc gcttctggga gttcgaagct ctgggctttc ttaacgaaga ccactggctt
 9181 ggaagaaaga actcaggagg cggggtcgag ggcttgggcc tccaaaaact gggttatatt
 9241 ctgcgtgaag ttggcacccg acctggaggc aagatctatg ctgatgacac agctggctgg
 9301 gacacccgca ttacgagagc tgacctggaa aatgaagcta aggttcttga gttgctggat
 9361 ggggaacatc ggcgtcttgc tagggccatc attgagctca cctatcgtca caaagttgtg
 9421 aaagtgatgc gcccggctgc tgatggaaga accgtcatgg atgtcatctc cagagaagat
 9481 cagaggggga gtggacaagt tgtcacctac gctctaaaca ccttcaccaa cctggccgtc
 9541 cagttggtga ggatgatgga aggggaagga gtgattggcc cagatgatgt ggagaaactc
 9601 acaaagggaa aaggacctaa agtcaggacc tggctgtttg agaatgggga ggaaagactc
 9661 agccgcatgg ctgtcagcgg agatgactgt gtggtaaagc ccctagatga ccgcttcgcc
 9721 acctctctcc acttcctcaa cgccatgtca aaggttcgca aagatatcca ggagtgggaaa
 9781 ccgtcaactg gatggtatga ctggcagcag gttccattct gctcgaacca tttcactgaa
 9841 ttaatcatga aagatggaag aacactggtg gttccatgcc gaggacagga cgaactggta
 9901 ggcagagctc gcatttctcc aggggccgga tggaacgtcc gtgacactgc ttgtctggct
 9961 aagtcttatg cccagatgtg gctgcttctg tacttccaca gaagagacct gcggctaatg
10021 gccaacgcca tttgctccgc tgtccctgtg aattgggtcc ctaccggaag aaccacgtgg
10081 tccatccatg ccggagggga gtggatgaca acagaagaca tgctggaggt ctggaaccgt
10141 gtttggatag aggagaatga atggatggaa gacaaaaccc cagtggagaa atggagtgac
10201 gtcccatact caggaaaacg ggaggacatc tggtgtggca gcttgattgg cacaagaacc
10261 cgagccacgt gggcagaaaa catccaggta gccatcaacc aagtcagagc aatcattgga
10321 gatgagaagt atgtggatta catgagttca ttaaagagat atgaagacac gactttggtt
10381 gaggacacag tactgtaaat actttattaa ttgtaaatag acaatgtaag catgtgtaaa
10441 agtatagttt tatagtagca tttagtgatg ttagtgtaaa tagttaagaa aattttaagg
10501 aggaagtcag gccggaaagt ttccgccacc ggaagttgag tagacggtgc tgcctgcgac
10561 tcaaccccag gaggactggg tgaacaaagc tgcgaagtga tccatgtaag ccctcagaac
10621 cgtctcggaa ggaggacccc acatgttgta acttcaaagc ccaatgtcag accacgctac
10681 ggcgtgccac tctgcggaga gtgcagtctg cgatagtgcc ccaggaggac tgggttaaca
10741 aaggcagatc aacgccccac gcggccctag ccctggtaat ggtgttaacc agggcgaaag
10801 gactagaggt tagaggagac cccgcggttt aaagtgcacg gcccagcctg actgaagctg
10861 taggtcaggg gaaggactag aggttagtgg agaccccgtg ccacaaaaca ccacaacaaa
10921 acagcatatt gacacctggg atagactagg agatcttctg ctctgcacaa ccagccacac
10981 ggcacagtgc gccgacaatg gtggctggtg gtgcgagaac acaggatct
```

//

Apr 17 2007 11:10:07

PubMed   Nucleotide   Protein   Genome   Structure   PMC   Taxonomy   OMIM   Books

Search : CoreNucleotide   for [   ] | Go | Clear |

Limits   Preview/Index   History   Clipboard   Details

Display GenBank   Show 5   Send to   Hide: ☐ sequence   ☐ all but gene, CDS and mRNA features

Range: from [begin]   to [end]   ☐ Reverse complemented strand   Features: [+] | Refresh |

☐ **1:** AF260969. Reports West Nile virus s...[gi:9930137]   Links

Features   Sequence

```
LOCUS       AF260969               11029 bp    RNA     linear   VRL 27-AUG-2000
DEFINITION  West Nile virus strain RO97-50, complete genome.
ACCESSION   AF260969
VERSION     AF260969.1  GI:9930137
KEYWORDS    .
SOURCE      West Nile virus
  ORGANISM  West Nile virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
            Flavivirus; Japanese encephalitis virus group.
REFERENCE   1  (bases 1 to 11029)
  AUTHORS   Savage,H.M., Ceianu,C., Nicolescu,G., Karabatsos,N., Lanciotti,R.,
            Vladimirescu,A., Laiv,L., Ungureanu,A., Romanca,C. and Tsai,T.F.
  TITLE     Entomologic and avian investigations of an epidemic of West Nile
            fever in Romania in 1996, with serologic and molecular
            characterization of a virus isolate from mosquitoes
  JOURNAL   Am. J. Trop. Med. Hyg. 61 (4), 600-611 (1999)
   PUBMED   10548295
REFERENCE   2  (bases 1 to 11029)
  AUTHORS   Bowen,M., Meyer,R.F., McKinney,N., Morrill,W. and Lanciotti,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-APR-2000) Arbovirus Diseases Branch, Centers for
            Disease Control & Prevention, Rampart Road, Fort Collins, CO 80521,
            USA
FEATURES             Location/Qualifiers
     source          1..11029
                     /organism="West Nile virus"
                     /mol_type="genomic RNA"
                     /strain="RO97-50"
                     /specific_host="Culex pipiens"
                     /db_xref="taxon:11082"
                     /country="Romania"
                     /note="1996"
     CDS             97..10398
                     /codon_start=1
                     /product="polyprotein precursor"
                     /protein_id="AAG02040.1"
                     /db_xref="GI:9930138"
                     /translation="MSKKPGGPGKSRAVNMLKRGMPRVLSLIGLKRAMLSLIDGKGPI
                     RFVLALLAFFRFTAIAPTRAVLDRWRGVNKQTAMKHLLSFKKELGTLTSAINRRSSKQ
                     KKRGGKTGIAVMIGLIASVGAVTLSNFQGKVMMTVNATDVTDVITIPTAAGKNLCIVR
                     AMDVGYMCDDTITYECPVLSAGNDPEDIDCWCTKSAVYVRYGRCTKTRHSRRSRRSLT
                     VQTHGESTLANKKGAWMDSTKATRYLVKTESWILRNPGYALVAAVIGWMLGSNTMQRV
                     VFVVLLLLVAPAYSFNCLGMSNRDFLEGVSGATWVDLVLEGDSCVTIMSKDKPTIDVK
                     MMNMEAANLAEVRSYCYLATVSDLSTKAACPTMGEAHNDKRADPAFVCRQGVVDRGWG
                     NGCGLFGKGSIDTCAKFACSTKATGRTILKENIKYEVAIFVHGPTTVESHGNYSTQIG
                     ATQAGRFSITPAAPSYTLKLGEYGEVTVDCEPRSGIDTNAYYVMTVGTKTFLVHREWF
                     MDLNLPWSSAGSTVWRNRETLMEFEEPHATKQSVIALGSQEGALHQALAGAIPVEFSS
                     NTVKLTSGHLKCRVKMEKLQLKGTTYGVCSKAFKFLGTPADTGHGTVVLELQYTGTDG
                     PCKVPISSVASLNDLTPVGRLVTVNPFVSVATANAKVLIELEPPFGDSYIVVGRGEQQ
```

```
                    INHHWHKSGSSIGKAFTTTLKGAQRLAALGDTAWDFGSVGGVFTSVGKAVHQVFGGAF
                    RSLFGGMSWITQGLLGALLLWMGINARDRSIALTFLAVGGVLLFLSVNVHADTGCAID
                    ISRQELRCGSGVFIHNDVEAWMDRYKYYPETPQGLAKIIQKAHKEGVCGLRSVSRLEH
                    QMWESVKDELNTLLKENGVDLSVVVEKQEGMYKSAPKRLTATTEKLEIGWKAWGKSIL
                    FAPELANNTFVVDGPETKECPTQNRAWNSLEVEDFGFGLTSTRMFLKVRESNTTECDS
                    KIIGTAVKNNLAIHSDLSYWIESRFNDTWKLERAVLGEVKSCTWPETHTLWGDGILES
                    DLIIPVTLAGPRSNHNRRPGYKTQNQGPWDEGRVEIDFDYCPGTTVTLSESCGHRGPA
                    TRTTTESGKLITDWCCRSCTLPPLRYQTDSGCWYGMEIRPQRHDEKTLVQSQVNAYNA
                    DMIDPFQLGLLVVFLATQEVLRKRWTAKISMPAILIALLVLVFGGITYTDVLRYVILV
                    GAAFAESNSGGDVVHLALMATFKIQPVFMVASFLKARWTNQENILLMLAAVFFQMAYH
                    DARQILLWEIPDVLNSLAVAWMILRAITFTTTSNVVVPLLALLTPGLRCLNLDVYRIL
                    LLMVGIGSLIREKRSAAAKKKGASLLCLALASTGLFNPMILAAGLIACDPNRKRGWPA
                    TEVMTAVGLMFAIVGGLAELDIDSMAIPMTIAGLMFAAFVISGKSTDMWIERTADISW
                    ESDAEITGSSERVDVRLDDDGNFQLMNDPGAPWKIWMLRMVCLAVSAYTPWAILPSVV
                    GFWITLQYTKRGGVLWDTPSPKEYKKGDTTTGVYRIMTRGLLGSYQAGAGVMVEGVFH
                    TLWHTTKGAALMSGEGRLDPYWGSVKEDRLCYGGPWKLQHKWNGQDEVQMIVVEPGKN
                    VKNVQTKPGVFKTPEGEIGAVTLDFPTGTSGSPIVDKNGDVIGLYGNGVIMPNGSYIS
                    AIVQGERMDEPIPAGFEPEMLRKKQITVLDLHPGAGKTRRILPQIIKEAINRRLRTAV
                    LAPTRVVAAEMAEALRGLPIRYQTSAVPREHNGNEIVDVMCHATLTHRLMSPHRVPNY
                    NLFVMDEAHFTDPASIAARGYISTKVELGEAAAIFMTATPPGTSDPFPESNSPISDLQ
                    TEIPDRAWNSGYEWITEYIGKTVWFVPSVKMGNEIALCLQRAGKKVVQLNRKSYETEY
                    PKCKNDDWDFVITTDISEMGANFKASRVIDSRKSVKPTIITEGEGRVILGEPSAVTAA
                    SAAQRRGRIGRNPSQVGDEYCYGGHTNEDDSNFAHWTEARIMLDNINMPNGLIAQFYQ
                    PEREKVYTMDGEYRLRGEERKNFLELLRTADLPVWLAYKVAAAGVSYHDRRWCFDGPR
                    TNTILEDNNEVEVITKLGERKILRPRWIDARVYSDHQALKAFKDFASGKRSQIGLIEV
                    LGKMPEHFMGKTWEALDTMYVVATAEKGGRAHRMALEELPDALQTIALIALLSVMTMG
                    VFFLLMQRKGIGKIGLGGVVLGVATFFCWMAEVPGTKIAGMLLLSLLLMIVLIPEPEK
                    QRSQTDNQLAVFLICVMTLVSAVAANEMGWLDKTKSDVSSLFGQRIEVKENFSMGEFL
                    LDLRPATAWSLYAVTTAVLTPLLKHLITSDYINTSLTSINVQASALFTLARGFPFVDV
                    GVSALLLAAGCWGQVTLTVTVTAATLLFCHYAYMVPGWQAEAMRSAQRRTAAGIMKNA
                    VVDGIVATDVPELERTTPIMQKKVGQIMLILVSLAAVVVNPSVKTVREAGILITAAAV
                    TLWENGASSVWNATTAIGLCHIMRGGWLSCLSITWTLMKNMDKPGLKRGGAKGRTLGE
                    VWKERLNQMTKEEFTRYRKEAIIEVDRSAAKHARKEGNVTGGHPVSRGTAKLRWLVER
                    RFLEPVGKVIDLGCGRGGWCYYMATQKRVQEVRGYTKGGPGHEEPQLVQSYGWNIVTM
                    KSGVDVFYRPSECCDTLLCDIGESSSSAEVEEHRTIRVLEMVEDWLHRGPREFCVKVL
                    CPYMPKVIEKMELLQRRYGGGLVRNPLSRNSTHEMYWVSRASGNVVHSVNMTSQVLLG
                    RMEKRTWKGPQYEEDVNLGSGTRAVGKPLLNSDTSKIKNRIERLRREYSSTWHHDENH
                    PYRTWNYHGSYDVKPTGSASSLVNGVVRLLSKPWDTITNVTTMAMTDTTPFGQQRVFK
                    EKVDTKAPEPPEGVKYVLNETTNWLWAFLAREKRPRMCSREEFIRKVNSNAALGAMFE
                    EQNQWRSAREAVEDPKFWEMVDEEREAHLRGECHTCIYNMMGKREKKPGEFGKAKGSR
                    AIWFMWLGARFLEFEALGFLNEDHWLGRKNSGGGVEGLGLQKLGYILREVGTRPGGKI
                    YADDTAGWDTRITRADLENEAKVLELLDGEHRRLARAIIELTYRHKVVKVMRPAADGR
                    TVMDVISREDQRGSGQVVTYALNTFTNLAVQLVRMMEGEGVIGPDDVEKLTKGKGPKV
                    RTWLFENGEERLSRMAVSGDDCVVKPLDDRFATSLHFLNAMSKVRKDIQEWKPSTGWY
                    DWQQVPFCSNHFTELIMKDGRTLVVPCRGQDELVGRARISPGAGWNVRDTACLAKSYA
                    QMWLLLYFHRRDLRLMANAICSAVPVNWVPTGRTTWSIHAGGEWMTTEDMLEVWNRVW
                    IEENEWMEDKTPVEKWSDVPYSGKREDIWCGSLIGTRARATWAENIQVAINQVRAIIG
                    DEKYVDYMSSLKRYEDTILVEDTVL"
mat_peptide        97..465
                   /product="nucleocapsid protein C"
mat_peptide        466..741
                   /product="pre-membrane protein prM"
mat_peptide        742..966
                   /product="membrane protein M"
mat_peptide        967..2469
                   /product="envelope glycoprotein E"
mat_peptide        2470..3525
                   /product="non-structural protein 1 NS1"
mat_peptide        3526..4218
                   /product="non-structural protein 2A NS2A"
mat_peptide        4219..4611
                   /product="non-structural protein 2B NS2B"
mat_peptide        4612..6468
                   /product="non-structural protein 3 NS3"
mat_peptide        6469..6915
                   /product="non-structural protein 4A NS4A"
```

```
     mat_peptide        6916..7680
                        /product="non-structural protein NS4B"
     mat_peptide        7681..10395
                        /product="non-structural protein NS5"
ORIGIN
        1 agtagttcgc ctgtgtgagc tgacaaactt agtagtgttt gtgaggatta acaacaatta
       61 acacagtgcg agctgtttct tagcacgaag atctcgatgt ctaagaaacc aggagggccc
      121 ggcaagagcc gggctgtcaa tatgctaaaa cgcggaatgc cccgcgtgtt gtccttgatt
      181 ggactgaaga gggcaatgtt gagcctgatt gacggcaagg ggccaatacg attcgtgttg
      241 gctctcttgg cgttcttcag gttcacagca attgctccga cccgagcagt actggaccga
      301 tggagaggtg tgaacaaaca aacagcgatg aaacaccttc tgagttttaa gaaagaacta
      361 gggaccttga ccagtgctat caatcggcgg agctcaaaac aaaagaaaag aggaggaaag
      421 accggaattg cagtcatgat tggcctgatc gccagcgtag gagcagttac cctctctaac
      481 ttccaaggga aagttatgat gacggtaaat gctactgacg tcacagatgt catcacgatt
      541 ccaacagctg ctggaaagaa cctatgcatt gtcagagcaa tggatgtggg atacatgtgt
      601 gatgacacta tcacttatga atgcccagtg ctgtcggccg gtaatgatcc agaagacatc
      661 gactgttggt gcacaaagtc agcagtctac gtcaggtatg gaagatgcac caagacacgc
      721 cactcaagac gcagtcggag gtcactgaca gtgcagacac acggagaaag cactctagcg
      781 aacaaaaagg gggcttggat ggacagcacc aaggccacaa ggtatctggt aaaaacagaa
      841 tcatggatct tgaggaaccc tggatatgcc ctggtggcag ctgttattgg ttggatgctt
      901 ggaagcaaca ccatgcagcg agttgtgttt gtagtgctat tgctcttggt ggctccagcc
      961 tacagcttca actgccttgg aatgagcaac agagacttct ggagggagt atctggagca
     1021 acttggggtgg atttggttct cgaaggcgac agctgcgtga ctatcatgtc taaggacaag
     1081 cctaccattg atgtgaagat gatgaatatg gaggctgcca acctggcaga ggtccgcagt
     1141 tattgctatt tggctaccgt cagtgatctc tccaccaaag ctgcatgccc gaccatggga
     1201 gaagctcaca atgacaaacg tgctgatcca gcttttgtgt gcagacaagg agtggtggac
     1261 aggggctggg gcaacggctg cgggctattt ggcaaaggaa gcattgacac atgcgccaaa
     1321 tttgcctgct ccaccaaggc aacaggaaga accatcttga aagaaatat caagtatgaa
     1381 gtggccatct ttgtccatgg accaaccact gtggagtcgc atggaaaacta ctccacacag
     1441 attgggggcca ctcaggcagg gagattcagc atcactcctg cggcgccttc atacacacta
     1501 aagcttggag aatatggaga agtgacagtg gactgtgaac cacggtcagg gattgacacc
     1561 aatgcttact acgtgatgac tgttggaaca aagacgtttt tggtccatcg tgagtggttc
     1621 atggacctca acctcccttg gagcagtgct ggaagtactg tgtggagaaa cagagagacg
     1681 ttaatggagt ttgaggaacc acacgccaca aagcagtctg tgatagcatt gggctcacaa
     1741 gagggagctc tgcatcaagc tttggctgga gccatccctg tggaattttc aagcaacact
     1801 gtcaagttga cgtcgggtca tttgaagtgt agagtgaaga tggaaaaatt gcagttgaag
     1861 ggaacaacct acggcgtctg ttcaaaggct ttcaagtttc ttgggactcc cgcagacaca
     1921 ggtcacggca ctgtggtgtt ggaattgcag tacactggca cggatggacc ttgcaaagtt
     1981 cccatctcgt cagtggcttc attgaacgac ctaacaccag tgggcagatt ggtcactgtc
     2041 aacccttttg tttcagtggc cacggccaat gccaaagtcc tgattgaact ggaaccaccc
     2101 tttggagact catacatagt ggtaggcaga ggagaacaac agatcaatca ccattggcat
     2161 aagtctggaa gcagcattgg caaagccttt acaaccaccc tcaaaggggc gcagagatta
     2221 gccgctctag agacacagac ttgggacttt ggatcagttg gagggggtgtt cacctcagta
     2281 gggaaggctg tccatcaagt gttcggtgga gcattccgct cactgttcgg aggtatgtcc
     2341 tggataacgc agggattgct gggggctctt ctgctgtgga tgggcatcaa tgcccgtgac
     2401 aggtccatag ctctcacgtt tctagcagtt ggaggagtcc tgctcttcct ctccgtgaac
     2461 gtgcacgccg acactgggtg tgccatagac atcagccggc aagagctgag atgtggaagt
     2521 ggagtgttca tacacaatga tgtggaggct tggatggacc gatacaagta ctaccctgaa
     2581 acgccacaag cctagccaa gatcattcaa aaagctcata aggaaggagt gtgcggtcta
     2641 cggtcagttt ccagactgga gcatcaaatg tgggaatcag tgaaggacga gctgaacact
     2701 ctttttgaaag agaatggtgt ggaccttagt gtcgtggttg agaaacagga gggaatgtac
     2761 aagtcggcac ctaaacgtct caccgccacc acggaaaaat tggaaattgg ctggaaggcc
     2821 tggggaaaga gcatcctatt tgcaccagaa cttgccaaca cacctttgt ggttgatggt
     2881 ccggagacca aggaatgccc aactcagaat cgcgcttgga atagcttgga agtggaggat
     2941 tttggatttg gtctcaccag cactcggatg ttcctgaagg tcagagagag caacacaact
     3001 gaatgtgact cggaagatca t cggaacggcc gtcaagaaca acttggcgat ccacagtgac
     3061 ctgtcctatt ggattgaaag caggttcaat gatacgtgga agcttgaaag ggcggttctg
     3121 ggtgaagtca aatcatgcac atggcctgag acgcacacct tgtggggcga tggaatcctc
     3181 gagagtgacc tgataatacc agtcacactg gcgggaccac gaagcaacca caatcggaga
     3241 cctgggtaca gacacaaaaa ccagggccca tgggacgaag tcgggtaga gattgatttc
     3301 gattactgcc cagggacac ggtcacccctg agtgagagct gcggacaccg tggacctgct
     3361 actcgcacca ccacagagag cggaaagttg ataacagatt ggtgctgcag gagctgcacc
     3421 ttaccaccat gcgctacca gacggacagc ggctgttggt atggcatgga gattagacca
     3481 cagagacatg atgaaaagac ccttgtgcag tcacaagtga atgcttacaa tgctgatatg
     3541 attgatccttt ttcagttggg ccttctggtc gtgttcttgg ccacccagga ggtccttcgc
     3601 aagaggtgga cagccaagat cagcatgcca gctatactga ttgctctgct agtcctggtg
```

```
3661 tttgggggca ttacctacac tgatgtgttg cgctatgtca tcttggtggg agcagctttc
3721 gcagaatcta attcgggagg agacgtggta cacttggcgc tcatggcgac ctttaagata
3781 cagccagtgt ttatggtggc atcgtttctc aaagcgagat ggaccaacca ggagaacatc
3841 ttgttgatgt tggcggctgt tttcttccaa atggcttacc acgatgcccg ccaaattctg
3901 ctctgggaga tccctgatgt gttgaattca ctggcagtag cttggatgat attgagagcc
3961 ataacattca caacgacatc aaacgtggtt gttccgttgc ttgccctgct aacacccggg
4021 ctgagatgct tgaatctgga tgtgtacagg atcctgctgt tgatggtcgg aataggcagt
4081 ttgatcaggg agaagagaag tgcagctgca aaaagaaag gagcaagtct gttatgcttg
4141 gctctagcct caacaggact tttcaatccc atgatccttg ccgctggact gattgcatgt
4201 gatcccaacc gtaaacgcgg gtggcccgca actgaagtga tgactgctgt cggcttgatg
4261 tttgccattg ttggagggct ggcagagctt gacattgact ccatggccat tccaatgact
4321 atcgcggggc tcatgtttgc tgctttcgtg atttctggga aatcaacaga tatgtggatc
4381 gagaggacgg cggacatctc ctgggaaagt gatgcagaaa tcacaggctc gagcgaaaga
4441 gttgatgtgc ggcttgatga tgatggaaac ttccagctca tgaatgatcc aggagcacct
4501 tggaaaatat ggatgctcag aatggtctgc ctcgcggtta gtgcgtacac cccttggccg
4561 atcctgccct cagtagttgg attttggata actctccagt acacaaagag aggaggcgtg
4621 ttgtgggaca ctccctcacc aaaggagtac aaaaaggggg acacgaccac cggagtctac
4681 aggatcatga ctcgtgggct gctcggcagt tatcaagcag gagcgggcgt gatggttgaa
4741 ggcgttttcc acactctttg gcacacaaca aaaggagccg ctttgatgag cggagaaggc
4801 cgcctggacc catactgggg tagtgtcaag gaggaccgac tttgttacgg aggaccctgg
4861 aaattgcagc ataagtggaa cgggcaggat gaggtgcaga tgattgtggt ggaacctggc
4921 aagaacgtca aaaacgtcca gacgaaacca ggggtgttca aaacacctga aggagaaatt
4981 ggggccgtga ctttggactt ccccactgga acatcaggct caccaatagt ggacaaaaac
5041 ggtgatgtga tcgggctcta tggcaatgga gtcataatgc ccaacggctc atacataagc
5101 gcgatagtgc agggtgaaag gatggatgag ccaatcccag ccggattcga acctgagatg
5161 ctgaggaaaa aacagatcac tgttctggat cttcatcccg gtgccggtaa aacaaggagg
5221 attctgccac agatcatcaa agaggccata aatagaagac tgagaacagc cgtgttagca
5281 ccaactaggg ttgtggctgc tgagatggct gaagccctga gggggactgcc tatccggtat
5341 cagacatctg cagtgcccag agaacacaat ggaaatgaga ttgttgatgt catgtcgcat
5401 gccaccctca ctcacaggct gatgtctcct cacagggtgc ctaactacaa tctgtttgtt
5461 atggatgagg ctcatttcac cgatccagct agcatcgcgg caagaggtta tatttccaca
5521 aaggtcgagc tgggggaggc ggcggcaata ttcatgacag ccaccccacc aggcacctca
5581 gatccattcc cagagtccaa ttcacctatt tctgacttac agactgagat cccggatcgg
5641 gcctggaact ctggatacga atggatcaca gaatacatcg ggaagacggt ttggtttgtg
5701 cctagtgtga aaatggggaa tgagattgcc ctttgtttac aacgtgccgg aaagaaagta
5761 gtccaattga acagaaagtc gtatgagacg gagtacccaa aatgtaagaa cgatgattgg
5821 gactttgtta tcacaacaga catatctgaa atgggggcta acttcaaggc gagcagggtg
5881 atcgacagcc ggaagagcgt gaagccaacc atcataacag aaggagaagg gagagtgatc
5941 ctgggagaac catctgcggt gacggcagct agtgccgccc agagacgtgg acgcatcggt
6001 agaaatccgt cgcaagttgg tgatgagtac tgttatggag ggcacacgaa tgaagacgac
6061 tcgaacttcg cccattggac tgaggcacga atcatgctgg acaacatcaa catgccaaac
6121 ggactgatcg ctcaattcta ccaaccagag cgtgagaagg tatacaccat ggatggggaa
6181 taccggctca gaggagaaga gaggaaaaac ttcctggaat tgttgaggac tgcagatctg
6241 ccagtttggc tggcttacaa ggtggcagcg gctggagtgt cataccacga tcggaggtgg
6301 tgttttgatg ccctaggac aaacacgatt ttagaagaca caacgaggt ggaagtcatc
6361 acgaagcttg gtgaaaggaa gattctgagg ccccgctgga ttgacgccag ggtgtactcg
6421 gatcaccagg cactaaaggc gttcaaggac tttgcttcgg gaaagcgttc tcagataggg
6481 ctcattgagg tcctgggaaa gatgcctgag catttcatgg ggaagacatg ggaagcactt
6541 gacaccatgt acgttgtggc cacagcagag aaaggaggaa gagctcacag aatggccctg
6601 gaggaactgc cggatgctct ccagacaatt gccttaattg ccttattgag tgtgatgacc
6661 atgggagtat tcttcctcct catgcagcgg aagggcattg gaaagatagg tttgggaggt
6721 gttgtcttgg gagtcgcgac cttcttctgt tggatggctg aagttccagg aacgaagatc
6781 gccggaatgt tgctgctttc ccttctcttg atgattgtgc taattcctga gccagagaag
6841 caacgttcgc agacagacaa ccagctagcc gtgttcctga tttgtgttat gacccttgtg
6901 agcgcagtgg cagccaacga aatgggttgg ctagataaga ccaagagtga cgtaagcagt
6961 ttgtttgggc aaagaattga ggtcaaggag aatttcagca tgggagagtt tcttctggac
7021 ttgagaccgg caacagcctg gtcactgtac gctgtgacca cagcggttct cactccactg
7081 ctaaagcatt tgatcacgtc agattacatt aacacttcat tgacctcaat taatgttcaa
7141 gcaagtgcac tattcacact cgcgcgaggt ttccccttcg tcgatgttgg agtgtcagct
7201 ctcttgctag cagccggatg ctgggacaa gtcacactca ccgtgacggt gacagcggca
7261 acactcctgt tctgtcacta tgcctacatg gttcccggat ggcaggctga ggcaatgcgc
7321 tcagcccagc ggcggacagc ggctgaatc atgaaaaacg ctgtagtgga tggcatcgtg
7381 gccacggacg tcccagaatt agagcgcacc acacccatca tgcagaagaa agtcggacaa
7441 atcatgctga tcttggtgtc tctagctgca gtagtagtga accgtctgt gaagacagtg
7501 cgagaagctg gaattttgat cacggcagca gcggtgacgc tttgggagaa tggagcaagc
7561 tctgtttgga acgcaacaac cgccatcgga ctctgccaca tcatgcgtgg gggttggttg
```

```
 7621 tcatgcctat ccataacatg gacactcatg aagaacatgg acaaaccagg actaaaaaga
 7681 ggtggggcaa aaggacgtac cttgggagag gtttggaaag aaagactcaa ccagatgaca
 7741 aaagaggagt tcactaggta ccgcaaagag gccatcatcg aagtcgaccg ctcagcagca
 7801 aaacacgcca ggaaagaggg caatgtcacc ggagggcatc cagtctctag aggcacagca
 7861 aaactgagat ggctggtcga acggaggttt ctcgaaccgg tcggaaaagt gattgacctt
 7921 ggatgtggaa gaggcggttg gtgttactat atggcaaccc aaaaaagagt ccaagaagtc
 7981 agagggtaca caaagggcgg tcccggacat gaagagcccc aactagtgca aagttatgga
 8041 tggaacattg tcaccatgaa gagtggagtg gatgtgttct atagaccttc tgagtgttgt
 8101 gacactctcc tttgtgacat cggagagtct tcgtcaagtg ctgaggttga agagcatagg
 8161 acgattcggg tccttgaaat ggttgaggac tggctgcacc gagggccaag ggaattttgt
 8221 gtgaaggtac tctgccccta catgccaaaa gtcatagaga agatggagct actccagcgc
 8281 cggtatgggg gaggactggt cagaaaccca ctctcacgga attccacgca cgagatgtac
 8341 tgggtaagtc gagcttcagg caatgtggta cattcagtga acatgactag ccaggtgctt
 8401 ctgggaagaa tggagaagag gacctggaag ggaccccaat acgaggaaga tgtaaacttg
 8461 ggaagtggaa ccagggcggt ggggaaaccc ctactcaact cagacaccag caaaatcaag
 8521 aacaggattg aacgactcag gcgtgagtac agttcaacgt ggcaccacga tgagaaccac
 8581 ccatatagaa cctggaacta tcacggcagt tatgacgtga agcccacagg ctccgctagc
 8641 tcgctggtta atggagtggt taggctcctc tcaaaaccat gggacaccat cacgaatgtc
 8701 accaccatgg ccatgactga cacaactccc ttcggacagc agcgggtgtt caaagagaag
 8761 gtggacacga aagctcctga accgccagaa ggagttaagt atgtgctcaa tgaaaccacc
 8821 aattggttgt gggcgtttct ggccagagaa aaacgtccca gaatgtgctc tcgagaggag
 8881 ttcataagaa aggtcaacag caacgcagct ttgggtgcca tgtttgaaga acagaatcaa
 8941 tggaggagcg ccagagaagc agttgaagat ccaaaatttt gggagatggt agatgaggaa
 9001 cgcgaggcac acctgcgggg ggaatgtcac acttgcatct acaacatgat gggaaagaga
 9061 gagaaaaaac ccggagagtt cggaaaggcc aagggaagca gagccatatg gttcatgtgg
 9121 ctcggagctc gctttctgga gttcgaagct ctgggttttc tcaatgaaga tcactggctt
 9181 ggaagaaaga actcaggagg aggtgtcgag ggcttgggtc tccaaaaact gggctacatc
 9241 ctgcgtgaag ttggcacccg acctgggggc aagatctatg ctgacgacac agctggctgg
 9301 gacacccgca tcacgagagc tgacttggaa aatgaagcta aggtgcttga gctgcttgat
 9361 ggggaacatc ggcgtcttgc cagggccatc attgagctca cctatcgtca caagttgtg
 9421 aaagtgatgc gcccggctgc tgatggaaga actgtcatgg atgtcatctc cagagaagat
 9481 cagaggggga gcggacaagt tgtcacctac gccttaaaca ctttcaccaa cctggccgtc
 9541 cagctggtga ggatgatgga aggggaagga gtgattggcc cagatgatgt ggagaaactc
 9601 acaaagggga aaggacccaa agtcaggacc tggctgtttg agaatggggga agagagactc
 9661 agccgcatgg ctgtcagcgg agatgactgt gtggtaaagc ctctggatga tcgctttgcc
 9721 acttcgctcc actttctcaa cgccatgtca aaggtccgca aagacattca agagtggaaa
 9781 ccatcaactg gatggtatga ttggcagcag gttccattct gctcaaacca tttcactgaa
 9841 ttgatcatga aagatggaag aacattggtg gttccatgcc gaggacagga cgaactggta
 9901 ggcagagctc gcatttctcc aggggccgga tggaacgtcc gcgacaccgc ttgtctggct
 9961 aagtcatatg cccagatgtg gctgcttctg tacttccaca gaagagacct gcggctcatg
10021 gccaacgcca tttgctccgc tgtccctgtg aactgggtcc ctaccggaag aaccacatgg
10081 tccatccatg caggaggaga gtggatgaca acagaggaca tgctggaggt ttggaaccgt
10141 gtttggatag aggagaatga atggatggaa gacaaaaccc cagtggagaa atggagtgac
10201 gtcccatact caggaaaacg ggaggacatc tggtgtggca gcctgattgg cacaagagcc
10261 cgagccacgt gggcagaaaa catccaggtg gctatcaacc aagtcagagc aattattgga
10321 gatgagaagt atgtggatta catgagttca ctaaagagat atgaagacac aattttggtt
10381 gaggacacag tactgtaaat actttatcaa ttgtaaatag ataatataag catgtatata
10441 ggtataggtt tatagtagca tttaatgatg ttagtgtaaa tagttaagaa aattttaagg
10501 aggaagtcag gccggaaaat ttccgccacc ggaagttgag tagacggtgc tgcctgcggc
10561 tcaacccccag gaggactggg tgaacaaagc cgcgaagtga tccatgtaag ccctcagaac
10621 cgtctcggaa ggaggacccc acatgttgta acttcaaagc ccaatgtcag accacgctac
10681 ggtgtgccac tctgcggaga gtgcagtctg cgatagtgcc ccaggaggac tgggttaaca
10741 aaggcaaacc aacgccccac gcggccctag ccctggtaat ggtgttaacc agggcgaaag
10801 gactagaggt tagaggagac cccgcggttt gaagtgcacg gcccagcctg gctgaagctg
10861 taggtcaggg gaaggactag aggttagtgg agaccccgtg ccacaaaaca ccacaacaaa
10921 acagcatatt gacacctggg atagactagg agatcttctg ctctgcacaa ccagccacac
10981 ggcacagtgc gccgacaatg gtggctggtg gtgcgagaac acaggatct
```

//

PubMed   Nucleotide   Protein   Genome   Structure   PMC   Taxonomy   OMIM   Books

My NCBI

[Sign In] [Register]

Nucleotide

Search   CoreNucleotide   for | Go | | Clear |

Limits   Preview/Index   History   Clipboard   Details

Display GenBank   Show 5   Send to   Hide: ☐ sequence ☐ all but gene, CDS and mRNA features

Range: from begin   to end   ☐ Reverse complemented strand   Features: |+| | Refresh |

☐ **1:** AF481864. Reports West Nile virus s...[gi:19387527]   Links

Features   Sequence

```
LOCUS       AF481864               11029 bp    RNA     linear   VRL 21-MAY-2002
DEFINITION  West Nile virus strain IS-98 STD, complete genome.
ACCESSION   AF481864
VERSION     AF481864.1  GI:19387527
KEYWORDS    .
SOURCE      West Nile virus
  ORGANISM  West Nile virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
            Flavivirus; Japanese encephalitis virus group.
REFERENCE   1  (bases 1 to 11029)
  AUTHORS   Malkinson,M., Banet,C., Weisman,Y., Pokamunski,S., King,R.,
            Drouet,M.T. and Deubel,V.
  TITLE     Introduction of West Nile virus in the Middle East by migrating
            white storks
  JOURNAL   Emerging Infect. Dis. 8 (4), 392-397 (2002)
   PUBMED   11971773
REFERENCE   2  (bases 1 to 11029)
  AUTHORS   Deubel,V., Malkinson,M. and Banet,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-FEB-2002) CERVI, Institut Pasteur, 21 Avenue Tony
            Garnier, Lyon 69365, France
FEATURES             Location/Qualifiers
     source          1..11029
                     /organism="West Nile virus"
                     /mol_type="genomic RNA"
                     /strain="IS-98 STD1"
                     /specific_host="sick stork"
                     /db_xref="taxon:11082"
                     /note="isolated in Israel in 1998"
     CDS             97..10398
                     /codon_start=1
                     /product="polyprotein precursor"
                     /protein_id="AAL87234.1"
                     /db_xref="GI:19387528"
                     /translation="MSKKPGGPGKSRAVNMLKRGMPRVLSLIGLKRAMLSLIDGKGPI
                     RFVLALLAFFRFTAIAPTRAVLDRWRGVNKQTAMKHLLSFKKELGTLTSAINRRSSKQ
                     KKRGGKTGIAVMIGLIASVGAVTLSNFQGKVMMTVNATDVTDVITIPTAAGKNLCIVR
                     AMDVGYMCDDTITYECPVLSAGNDPEDIDCWCTKSAVYVRYGRCTKTRHSRRSRRSLT
                     VQTHGESTLANKKGAWMDSTKATRYLVKTESWILRNPGYALVAAVIGWMLGSNTMQRV
                     VFVVLLLLVAPAYSFNCLGMSNRDFLEGVSGATWVDLVLEGDSCVTIMSKDKPTIDVK
                     MMNMEAANLAEVRSYCYLATVSDLSTKAACPTMGEAHNDKRADPAFVCRQGVVDRGWG
                     NGCGLFGKGSIDTCAKFACSTKAIGRTILKENIKYEVAIFVHGPTTVESHGNYSTQVG
                     ATQAGRFSITPAAPSYTLKLGEYGEVTVDCEPRSGIDTNAYYVMTVGTKTFLVHREWF
                     MDLNLPWSSAGSTVWRNRETLMEFEEPHATKQSVIALGSQEGALHQALAGAIPVEFSS
                     NTVKLTSGHLKCRVKMEKLQLKGTTYGVCSKAFKFLGTPADTGHGTVVLELQYTGTDG
                     PCKVPISSVASLNDLTPVGRLVTVNPFVSVATANAKVLIELEPPFGDSYIVVGRGEQQ
                     INHHWHKSGSSIGKAFTTTLKGAQRLAALGDTAWDFGSVGGVFTSVGKAVHQVFGGAF
                     RSLFGGMSWITQGLLGALLLWMGINARDRSIALTFLAVGGVLLFLSVNVHADTGCAID
                     ISRQELRCGNGVFIHNDVEAWMDRYKYYPETPQGLAKIIQKAHKEGVCGLRSVSRLEH
```

```
QMWEAVKDELNTLLKENGVDLSVVVEKQEGMYKSAPKRLTATTEKLEIGWKAWGKSIL
FAPELANNTFVVDGPETKECPTQNRAWNSLEVEDFGFGLTSTRMFLKVRESNTTECDS
KIIGTAVKNNLAIHSDLSYWIESRLNDTWKLERAVLGEVKSCTWPETHTLWGDGILES
DLIIPVTLAGPRSNHNRRPGYKTQNQGPWDEGRVEIDFDYCPGTTVTLSESCGHRGPA
TRTTTESGKLITDWCCRSCTLPPLRYQTDSGCWYGMEIRPQRHDEKTLVQSQVNAYNA
DMIDPFQLGLLVVFLATQEVLRKRWTAKISMPAILIALLVLVFGGITYTDVLRYVILV
GAAFAESNSGGDVVHLALMATFKIQPVFMVASFLKARWTNQENILLMLAAVFFQMAYH
DARQILLWEIPDVLNSLAVAWMILRAITFTTTSNVVVPLLALLTPRLRCLNLDVYRIL
LLMVGIGSLIREKRSAAAKKKGASLLCLALASTGLFNPMILAAGLIACDPNRKRGWPA
TEVMTAVGLMFAIVGGLAELDIDSMAIPMTIAGLMFAAFVISGKSTDMWIERTADISW
ESDAEITGSSERVDVRLDDGENFQLMNDPGAPWKIWMLRMVCLAISAYTPWAILPSVV
GFWITLQYTKRGGVLWDTPSPKEYKKGDTTTGVYRIMTRGLLGSYQAGAGVMVEGVFH
TLWHTTKGAALMSGEGRLDPYWGSVKEDRLCYGGPWKLQHKWNGQDEVQMIVVEPGKN
VKNVQTKPGVFKTPEGEIGAVTLDFPTGTSGSPIVDKNGDVIGLYGNGVIMPNGSYIS
AIVQGERMDEPIPAGFEPEMLRKKQITVLDLHPGAGKTRRILPQIIKEAINRRLRTAV
LAPTRVVAAEMAEALRGLPIRYQTSAVPREHNGNEIVDVMCHATLTHRLMSPHRVPNY
NLFVMDEAHFTDPASIAARGYISTKVELGEAAAIFMTATPPGTSDPFPESNSPISDLQ
TEIPDRAWNSGYEWITEYTGKTVWFVPSVKMGNEIALCLQRAGKKVVQLNRKSYETEY
PKCKNDDWDFVITTDISEMGANFKASRVIDSRKSVKPTIITEGEGRVILGEPSAVTAA
SAAQRRGRIGRNPSQVGDEYCYGGHTNEDDSNFAHWTEARIMPDNINMPNGLIAQFYQ
PEREKVYTMEGEYRLRGEERKNFLELLRTADLPVWLAYKVAAAGVSYHDRRWCFDGPR
TNTILEDNNEVEVITKLGERKILRPRWIDARVYSDHQALKAFKDFASGKRSQIGLIEV
LGKMPEHFMGKTWEALDTMYVVATAEKGGRAHRMALEELPDALQTIALIALLSVMTMG
VFFLLMQRKGIGKIGLGGAVLGVATFFCWMAEVPGTKIAGMLLLSLLLMIVLIPEPEK
QRSQTDNQLAVFLICVMTLVSAVAANEMGWLDKTKSDISSLFGQRIEVKENFSMGEFL
LDLRPATAWSLYAVTTAVLTPLLKHLITSDYINTSLTSINVQASALFTLARGFPFVDV
GVSALLLAAGCWGQVTLTVTVTAATLLFCHYAYMVPGWQAEAMRSAQRRTAAGIMKNA
VVDGIVATDVPELERTTPIMQKKVGQIMLILVSLAAVVVNPSVKTVREAGILITAAAV
TLWENGASSVWNATTAIGLCHIMRGGWLSCLSITWTLIKNMEKPGLKRGGAKGRTLGE
VWKERLNQMTKEEFTRYRKEAIIEVDRSAAKHARKEGNVTGGHSVSRGTAKLRWLVER
RFLEPVGKVIDLGCGRGGWCYYMATQKRVQEVRGYTKGGPGHEEPQLVQSYGWNIVTM
KSGVDVFYRPSECCDTLLCDIGESSSSAEVEEHRTIRVLEMVEDWLHRGPREFCVKVL
CPYMPKVIEKMELLQRRYGGGLVRNPLSRNSTHEMYWVSRASGNVVHSVNMTSQVLLG
RMEKRTWKGPQYEEDVNLGSGTRAVGKPLLNSDTSKINNRIERLRREYSSTWHHDENH
PYRTWNYHGSYDVKPTGSASSLVNGVVRLLSKPWDTITNVTTMAMTDTTPFGQQRVFK
EKVDTKAPEPPEGAKYVLNETTNWLWAFLAREKRPRMCSREEFIRKVNSNAALGAMFE
EQNQWRSAREAVEDPKFWEMVDEEREAHLRGECHTCIYNMMGKREKKPGEFGKAKGSR
AIWFMWLGARFLEFEALGFLNEDHWLGRKNSGGGVEGLGLQKLGYILREVGTRPGGKI
YADDTAGWDTRITRADLENEAKVLELLDGEHRRLARAIIELTYRHKVVKVMRPAADGR
TVMDVISREDQRGSGQVVTYALNTFTNLAVQLVRMMEGEGVIGPDDVEKLTKGKGPKV
RTWLFENGEERLSRMAVSGDDCVVKPLDDRFATSLHFLNAMSKVRKDIQEWKPSTGWY
DWQQVPFCSNHFTELIMKDGRTLVVPCRGQDELVGRARISPGAGWNVRDTACLAKSYA
QMWLLLYFHRRDLRLMANAICSAVPVNWVPTGRTTWSIHAGGEWMTTEDMLEVWNRVW
IEENEWMEDKTPVEKWSDVPYSGKREDIWCGSLIGTRARATWAENIQVAINQVRAIIG
DEKYVDYMSSLKRYEDTTLVEDTVL"
```

ORIGIN

```
   1 agtagttcgc ctgtgtgagc tgacaaactt agtagtgttt gtgaggatta acaacaatta
  61 acacagtgcg agctgtttct tagcacgaag atctcgatgt ctaagaaacc aggagggccc
 121 ggcaagagcc gggctgtcaa tatgctaaaa cgcggaatgc cccgcgtgtt gtccttgatt
 181 ggactgaaga gggctatgtt gagcctgatc gacggcaagg ggccaatacg atttgtgttg
 241 gctctcttgg cgttcttcag gttcacagca attgctccga cccgagcagt gctggatcga
 301 tggagaggtg tgaataaaca aacagcgatg aaacaccttc tgagttttaa gaaggaacta
 361 gggaccttga ccagtgctat caatcggcgg agctcaaaac aaaagaaaag aggaggaaag
 421 accggaattg cagtcatgat tggcctgatc gccagcgtag gagcagttac cctctctaac
 481 ttccaaggga aggtgatgat gacggtaaat gctactgacg tcacagatgt catcacgatt
 541 ccaacagctg ctggaaagaa cctatgcatt gtcacagcaa tggatgtgaa atacagtgtgc
 601 gatgatacta tcacttatga atgcccagtg ctgtcggctg gtaatgatcc agaagacatc
 661 gactgttggt gcacaaagtc agcagtctac gtcaggtatg gaagatgcac caagacacgc
 721 cactcaagac gcagtcggag gtcactgaca gtgcagacac acggagaaag cactctagcg
 781 aacaagaagg gggcttggat ggacagcacc aaggccacaa ggtatttggt aaaaacagaa
 841 tcatggatct tgaggaaccc tggatatgcc ctggtggcag ccgtcattgg ttggatgctt
 901 gggagcaaca ccatgcagag agttgtgttt gtcgtgctat tgcttttggt ggccccagct
 961 tacagcttta actgccttgg aatgagcaac agagacttct tggaaggagt gtctggagca
1021 acatgggtgg atttggttct cgaaggcgac agctgcgtga ctatcatgtc taaggacaag
1081 cctaccatcg atgtgaagat gatgaatatg gaggcggcca acctggcaga ggtccgcagt
1141 tattgctatt tggctaccgt cagcgatctc tccaccaaag ctgcgtgccc gaccatggga
```

```
1201 gaagctcaca atgacaaacg tgctgaccca gcttttgtgt gcagacaagg agtggtggac
1261 aggggctggg gcaacggctg cggactattt ggcaaaggaa gcattgacac atgcgccaaa
1321 tttgcctgct ctaccaaggc aataggaaga accatcttga aagagaatat caagtacgaa
1381 gtggccattt ttgtccatgg accaactact gtggagtcgc acggaaacta ctccacacag
1441 gttggagcca ctcaggcagg gagattcagc atcactcctg cggcgccttc atacacacta
1501 aagcttggag aatatggaga ggtgacagtg gactgtgaac cacggtcagg gattgacacc
1561 aatgcatact acgtgatgac tgttggaaca aagacgttct tggtccatcg tgagtggttc
1621 atggacctca acctcccttg gagcagtgct ggaagtactg tgtggaggaa cagagagacg
1681 ttaatggagt ttgaggaacc acacgccacg aagcagtctg tgatagcatt gggctcacaa
1741 gagggagctc tgcatcaagc tttggctgga gccattcctg tggaattttc aagcaacact
1801 gtcaagttga cgtcgggtca tttgaagtgt agagtgaaga tggaaaaatt gcagttgaag
1861 ggaacaacct atggcgtctg ttcaaaggct ttcaagtttc ttgggactcc cgcagacaca
1921 ggtcacggca ctgtggtgtt ggaattgcag tacactggca cggatggacc ttgcaaagtt
1981 cctatctcgt cagtggcttc attgaacgac ctaacgccag tgggcagatt ggtcactgtc
2041 aacccttttg tttcagtggc cacggccaac gctaaggtcc tgattgaatt ggaaccaccc
2101 tttggagact catacatagt ggtgggcaga ggagaacaac agattaatca ccattggcac
2161 aagtctggaa gcagcattgg caaagccttt acaaccaccc tcaaaggagc gcagagacta
2221 gccgctctag agacacagc ttgggacttt ggatcagttg gaggggtgtt cacctcagtt
2281 gggaaggctg tccatcaagt gttcggagga gcattccgct cactgttcgg aggcatgtcc
2341 tggataacgc aaggattgct gggggctctc ctgttgtgga tgggcatcaa tgctcgtgat
2401 aggtccatag ctctcacgtt tctcgcagtt ggaggagttc tgctcttcct ctccgtgaac
2461 gtgcacgctg acactgggtg tgccatagac atcagccggc aagagctgag atgtggaaat
2521 ggagtgttca tacacaatga tgtggaggct tggatggacc ggtacaagta ttaccctgaa
2581 acgccacaag gcctagccaa gatcattcag aaagctcata aggaaggagt gtgcggtcta
2641 cgatcagttt ccagactgga gcatcaaatg tgggaagcag tgaaggacga gctgaacact
2701 cttttgaagg agaatggtgt ggaccttagt gtcgtggttg agaaacagga gggaatgtac
2761 aagtcagcac ctaaacgcct caccgccacc acggaaaaat tggaaattgg ctggaaggcc
2821 tggggaaaga gtattttatt tgcaccagaa ctcgccaaca acacctttgt ggttgatggt
2881 ccggagacca aggaatgtcc gactcagaat cgcgcttgga atagcttaga agtggaggat
2941 tttggatttg gtctcaccag cactcggatg ttcctgaagg tcagagagag caacacaact
3001 gaatgtgact cgaagatcat tggaacggct gtcaagaaca acttggcgat ccacagtgac
3061 ctgtcctatt ggattgaaag caggctcaat gatacgtgga gcttgaaag ggcagttctg
3121 ggtgaagtca aatcatgtac gtggcctgag acgcatacct tgtggggcga tggaatcctt
3181 gagagtgact tgataatacc agtcacactg gcgggaccac gaagcaatca caatcggaga
3241 cctgggtaca agacacaaaa ccagggccca tgggacgaag gccgggtaga gattgacttc
3301 gattactgcc caggaactac ggtcaccctg agtgagagct gcggacaccg tggacctgcc
3361 actcgcacca ccacagagag cggaaagttg ataacagatt ggtgctgcag gagctgcacc
3421 ttaccaccac tgcgctacca aactgacagc ggctgttggt atggtatgga gatcagacca
3481 cagagacatg atgaaaagac cctcgtgcag tcacaagtga atgcttataa tgctgatatg
3541 attgaccctt ttcagttggg ccttctggtc gtgttcttgg ccacccagga ggtccttcgc
3601 aagaggtgga cagccaagat cagcatgcca gctatactga ttgctctgct agtcctggtg
3661 tttgggggca ttacttacac tgatgtgtta cgctatgtca tcttggtggg ggcagctttc
3721 gcagaatcta attcgggagg agacgtggta cacttggcgc tcatggcgac cttcaagata
3781 caaccagtgt ttatggtggc atcgtttctc aaagcgagat ggaccaacca ggagaacatt
3841 ttgttgatgt tggcggctgt tttctttcaa atggcttatc acgatgcccg ccaaattctg
3901 ctctgggaga tccctgatgt gttgaattca ctggcggtag cttggatgat actgagagcc
3961 ataacattca caacgacatc aaatgtggtc gtcccgctgc tagccctgct aacacccggg
4021 ctgagatgct tgaatctgga tgtgtacagg atactgctgt tgatggtcgg aataggcagc
4081 ttgatcaggg agaagaggag tgcagctgca aaaaagaaag gagcaagtct gctatgcttg
4141 gctctagcct caacaggact tttcaacccc atgatccttg ctgctggact gattgcatgt
4201 gatcccaacc gtaaacgcgg atggccgca actgaagtga tgacagctgt cggcctaatg
4261 tttgccatcg tcggagggct ggcagagctt gacattgact ccatggccat tccaatgact
4321 atcgcggggc tcatgtttgc tgctttcgtg atttctggga aatcaacaga tatgtggatt
4381 gagagaacgg cggacatttc ctgggaaagt gatgcagaaa ttacaggctc gagcgaaaga
4441 gttgatgttc ggcttgatga tggtgaaaac ttccagctca tgaatgatcc aggagcacct
4501 tggaagatat ggatgctcag aatggtctgt ctcgcgatta gtgcgtacac ccctgggca
4561 atcttgccct cagtagttgg attttggata actctccaat acacaaagag aggaggtgtg
4621 ttgtgggaca ctccctcacc aaaggagtac aaaaagggg acacgaccac cggcgtctac
4681 aggatcatga ctcgtgggct gctcggcagt tatcaagcag gagcgggcgt gatggttgaa
4741 ggtgtttttc acacccttg gcatacaaca aaaggagccg ctttgatgag cggagagggc
4801 cgcctggacc catactgggg cagtgtcaag gaggatcgac tttgttacgg aggaccctgg
4861 aaattgcagc acaagtggaa cgggcaggat gaggtgcaga tgattgtggt ggaacctggc
4921 aagaacgtta agaacgtcca gacgaaacca ggggtgttca aaacacctga aggagaaatc
4981 ggggccgtga ctttggactt ccccactgga acatcaggct caccaatagt ggacaaaaac
5041 ggtgatgtga ttgggcttta tggcaatgga gtcataatgc ccaacggctc atacataagc
5101 gcgatagtgc agggtgaaag gatggatgag ccaatcccag ccggattcga acctgagatg
```

```
5161 ctgaggaaaa aacagatcac tgtactggat ctccatcccg gcgccggtaa aacaaggagg
5221 attctgccac agatcatcaa agaggccata aacagaagac tgagaacagc cgtgctagca
5281 ccaaccaggg ttgtggctgc tgagatggct gaagcactga gaggactgcc catccggtac
5341 cagacatccg cagtgcccag agaacataat ggaaatgaga ttgttgatgt catgtgtcat
5401 gctaccctca cccacaggct gatgtctcct cacagggtgc cgaactacaa cctgttcgtg
5461 atggatgagg ctcatttcac cgacccagct agtatcgcag caagaggtta catttccaca
5521 aaggtcgagc taggggaggc ggcggcaata ttcatgacag ccacecccacc aggcacttca
5581 gatccattcc cagagtccaa ttcaccaatt tccgacttac agactgagat cccggatcga
5641 gcttggaact ctggatacga atggatcaca gaatacaccg gaagacggt ttggtttgtg
5701 cctagtgtca agatggggaa tgagattgcc ctttgcctac aacgtgctgg aaagaaagta
5761 gtccaattga acagaaagtc gtacgagacg gagtacccaa aatgtaagaa cgatgattgg
5821 gactttgtta tcacaacaga catatctgaa atggggggcta acttcaaggc gagcagggtg
5881 attgacagcc ggaagagtgt gaaaccaacc atcataacag aaggagaagg gagagtgatc
5941 ctgggagaac catctgcagt gacagcagct agtgccgccc agagacgtgg acgtatcggt
6001 agaaatccgt cgcaagttgg tgatgagtac tgttatgggg ggcacacgaa tgaagacgac
6061 tcgaacttcg cccattggac tgaggcacga atcatgccgg acaacatcaa catgccaaac
6121 ggactgatcg ctcaattcta ccaaccagag cgtgagaagg tatataccat ggaggggggaa
6181 taccggctca gaggagaaga gaggaaaaac tttctggaac tgttgaggac tgcagatctg
6241 ccagtttggc tggcttacaa ggttgcagcg gctggagtgt cataccacga ccggaggtgg
6301 tgctttgatg gtcctaggac aaacacaatt ttagaagaca caacgaagt ggaagtcatc
6361 acgaagcttg gtgaaaggaa gattctgagg ccgcgctgga ttgacgccag ggtgtactcg
6421 gatcaccagg cactaaaggc gttcaaggac ttcgcctcgg gaaaacgttc tcagataggg
6481 ctcattgagg ttctgggaaa gatgcctgag cacttcatgg ggaagacatg ggaagcactt
6541 gacaccatgt acgttgtggc cactgcagag aaaggaggaa gagctcacag aatggccctg
6601 gaggaactgc cagatgctct tcagacaatt gccttgattg ccttattgag tgtgatgacc
6661 atgggagtat tcttcctcct catgcagcgg aagggcattg gaaagatagg tttgggaggc
6721 gctgtcttgg gagtcgcgac ctttttctgt tggatggctg aagttccagg aacgaagatc
6781 gccggaatgt tgctgctctc ccttctcttg atgattggtc taattcctga gccagagaag
6841 caacgttcgc agacagacaa ccagctagcc gtgttcctga tttgtgtcat gacccttgtg
6901 agcgcagtgg cagccaacga gatgggttgg ctagacaaga ccaagagtga cataagcagt
6961 ttgtttgggc aaagaattga ggtcaaggag aatttcagca tgggagagtt tcttctggac
7021 ttgaggccgg caacagcctg gtcactgtac gctgtgacaa cagcggtcct cactccactg
7081 ctaaagcatt tgatcacgtc agattacatc aacacctcat tgacctcaat aaacgttcag
7141 gcaagtgcac tattcacact cgcgcgaggc ttccccttcg tcgatgttgg agtgtcggct
7201 ctcctgctag cagccggatg ctggggacaa gtcaccctca ccgttacggt aacagcggca
7261 acactccttt tttgccacta tgcctacatg gttcccggtt ggcaagctga ggcaatgcgc
7321 tcagcccagc ggcggacagc ggccggaatc atgaaaaacg ctgtagtgga tggcatcgtg
7381 gccacggacg tcccagaatt agagcgcacc acaccatca tgcagaagaa agttggacag
7441 atcatgctga tcttggtgtc tctagctgca gtagtagtga acccgtctgt gaagacagta
7501 cgagaagccg gaattttgat cacggccgca gcggtgacgc tttgggagaa tggagcaagc
7561 tctgtttgga acgcaacaac tgccatcgga ctctgccaca tcatgcgtgg gggttggttg
7621 tcatgtctat ccataacatg gacactcata aagaacatgg aaaaaccagg actaaaaaga
7681 ggtggggcaa aaggacgcac cttgggagag gtttggaaag aaagactcaa ccagatgaca
7741 aaagaagagt tcactaggta ccgcaaagag gccatcatcg aagtcgatcg ctcagcggca
7801 aaacacgcca ggaaagaagg caatgtcact ggagggcatt cagtctctag gggcacagca
7861 aaactgagat ggctggtcga acggaggttt ctcgaaccgg tcggaaaagt gattgacctt
7921 ggatgtggaa gaggcggttg gtgttactat atggcaaccc aaaaaagagt ccaagaagtc
7981 agagggtaca caaagggcgg tcccggacat gaagagcccc aactagtgca aagttatgga
8041 tggaacattg tcaccatgaa gagtggagtg gatgtgttct acagaccttc tgagtgttgt
8101 gacaccctcc tttgtgacat cggagagtcc tcgtcaagtc tgaggttga agagcatagg
8161 acgattcgtg tccttgaaat ggttgaggac tggctgcacc gagggccaag ggaattttgc
8221 gtgaaggtgc tctgccccta catgccgaaa gtcatagaga agatggagct gctccaacgc
8281 cggtatgggg ggggactggt cagaaaccca ctctcacgga attccacgca cgagatgtat
8341 tgggtgagtc gagcttcagg caatgtggta cattcagtga atatgaccag ccaggtgctc
8401 ctaggaagaa tggaaaaaga gacctggaag ggaccccaat acgaggaaga tgtaaacttg
8461 ggaagcggaa ccagggcggt gggaaaaacc ctgctcaact cagacaccag taaaatcaac
8521 aacaggattg aacgactcag gcgtgagtac agttcgacgt ggcaccacga tgagaaccac
8581 ccatatagaa cctggaacta tcacggcagt tatgatgtga gcccacagg ctccgccagt
8641 tcgctggtca atggagtggt caggctcctc tcaaaaccat gggacaccat cacgaatgtt
8701 accaccatgg ccatgactga cactactccc ttcgggcagc agcgagtgtt caaagagaag
8761 gtggacacga aagctcctga accgccagaa ggagcgaagt acgtgctcaa cgagaccacc
8821 aactggttgt gggcgttttt ggccagagaa aaacgtccca gaatgtgctc tcgagaggaa
8881 ttcataagaa aggtcaacag caatgcagct ttgggtgcca tgtttgaaga gcagaatcaa
8941 tggaggagcg ccagagaagc agttgaagat ccaaaatttt gggagatggt ggatgaggag
9001 cgcgaggcac atctgcgggg ggaatgtcac acttgcattt acaacatgat gggaaagaga
9061 gagaaaaaac ccggagagtt cggaaaggcc aagggaagca gagccatttg gttcatgtgg
```

```
 9121 ctcggagctc gctttctgga gttcgaggct ctgggttttc tcaatgaaga ccactggctt
 9181 ggaagaaaga actcaggagg aggtgtcgag ggcttgggcc tccaaaaact gggttacatc
 9241 ctgcgtgaag ttggcacccg gcctgggggc aagatctatg ctgatgacac agctggctgg
 9301 gacacccgca tcacgagagc tgacttggaa aatgaagcta aggtgcttga gctgcttgat
 9361 ggggaacatc ggcgtcttgc cagggccatc attgagctca cctatcgtca caaagttgtg
 9421 aaagtgatgc gcccggctgc tgatggaaga accgtcatgg atgttatctc cagagaagat
 9481 cagaggggga gtggacaagt tgtcacctac gccctaaaca ctttcaccaa cctggccgtc
 9541 cagctggtga ggatgatgga aggggaagga gtgattggcc cagatgatgt ggagaaactc
 9601 acaaaaggga aaggacccaa agtcaggacc tggctgtttg agaatgggga agaaagactc
 9661 agccgcatgg ctgtcagtgg agatgactgt gtggtaaagc ccctggacga tcgctttgcc
 9721 acctcgctcc acttcctcaa tgctatgtca aaggttcgca aagacatcca agagtggaaa
 9781 ccgtcaactg gatggtatga ttggcagcag gttccatttt gctcaaacca tttcactgaa
 9841 ttgatcatga aagatggaag aacactggtc gttccatgcc gaggacagga tgaattggta
 9901 ggcagagctc gcatatctcc aggggccgga tggaacgtcc gcgacactgc ttgtctggct
 9961 aagtcttatg cccagatgtg gctgcttctg tacttccaca gaagagacct gcggctcatg
10021 gccaacgcca tttgctccgc tgtccctgtg aattgggtcc ctaccggaag aaccacgtgg
10081 tccatccatg caggaggaga gtggatgaca acagaggaca tgttggaggt ctggaaccgt
10141 gtttggatag aggagaatga atggatggaa gacaaaaccc cagtggagaa atggagtgac
10201 gtcccatatt caggaaaacg agaggacatc tggtgtggca gcctgattgg cacaagagcc
10261 cgagccacgt gggcagaaaa catccaggtg gctatcaacc aagtcagagc aatcatcgga
10321 gatgagaagt atgtggatta catgagttca ctaaagagat atgaagacac aactttggtt
10381 gaggacacag tactgtagat atttaatcaa ttgtaaatag acaatataag tatgcataaa
10441 agtgtagttt tatagtagta tttagtggtg ttagtgtaaa tagttaagaa aattttgagg
10501 agaaagtcag gccgggaagt tcccgccacc ggaagttgag tagacggtgc tgcctgcgac
10561 tcaaccccag gaggactggg tgaacaaagc cgcgaagtga tccatgtaag ccctcagaac
10621 cgtctcggaa ggaggacccc acatgttgta acttcaaagc ccaatgtcag accacgctac
10681 ggcgtgctac tctgcggaga gtgcagtctg cgatagtgcc ccaggaggac tgggttaaca
10741 aaggcaaacc aacgccccac gcggccctag ccccggtaat ggtgttaacc agggcgaaag
10801 gactagaggt tagaggagac cccgcggttt aaagtgcacg gcccagcctg actgaagctg
10861 taggtcaggg gaaggactag aggttagtgg agaccccgtg ccacaaaaca ccacaacaaa
10921 acagcatatt gacacctggg atagactagg agatcttctg ctctgcacaa ccagccacac
10981 ggcacagtgc gccgacaatg gtggctggtg gtgcgagaac acaggatct
//
```

NCBI — Nucleotide

PubMed   Nucleotide   Protein   Genome   Structure   PMC   Taxonomy   OMIM   Books

Search   CoreNucleotide   for   [   ]   [Go] [Clear]

Limits   Preview/Index   History   Clipboard   Details

Display GenBank   Show 5   Send to   Hide: ☐ sequence ☐ all but gene, CDS and mRNA features

Range: from ˙begin   to ˙end   ☐ Reverse complemented strand   Features: [+] [Refresh]

☐1: AF404753. Reports West Nile virus i...[gi:21929232]   Links

Features   Sequence

```
LOCUS       AF404753               11029 bp ss-RNA     linear   VRL 23-JUL-2002
DEFINITION  West Nile virus isolate WN MD 2000-crow265, complete genome.
ACCESSION   AF404753
VERSION     AF404753.1  GI:21929232
KEYWORDS    .
SOURCE      West Nile virus
  ORGANISM  West Nile virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
            Flavivirus; Japanese encephalitis virus group.
REFERENCE   1  (bases 1 to 11029)
  AUTHORS   Lanciotti,R.S., Ebel,G.D., Deubel,V., Kerst,A.J., Murri,S.,
            Meyer,R., Bowen,M., McKinney,N., Morrill,W.E., Crabtree,M.B.,
            Kramer,L.D. and Roehrig,J.T.
  TITLE     Complete genome sequences and phylogenetic analysis of West Nile
            virus strains isolated from the United States, Europe, and the
            Middle East
  JOURNAL   Virology 298 (1), 96-105 (2002)
   PUBMED   12093177
REFERENCE   2  (bases 1 to 11029)
  AUTHORS   Lanciotti,R.S., Ebel,G.D. and Kerst,A.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Division of Vector-Borne Infectious
            Diseases, Centers for Disease Control & Prevention, Rampart Road,
            Fort Collins, CO 80521, USA
FEATURES             Location/Qualifiers
     source          1..11029
                     /organism="West Nile virus"
                     /mol_type="genomic RNA"
                     /isolate="WN MD 2000-crow265"
                     /specific_host="crow"
                     /db_xref="taxon:11082"
                     /country="USA: Maryland"
     CDS             97..10398
                     /note="contains capsid, pre-membrane, envelope, NS1, NS2a,
                     NS2b, NS3, NS4a, NS4b, and NS5"
                     /codon_start=1
                     /product="polyprotein precursor"
                     /protein_id="AAM81749.1"
                     /db_xref="GI:21929233"
                     /translation="MSKKPGGPGKSRAVNMLKRGMPRVLSLIGLKRAMLSLIDKGPI
                     RFVLALLAFFRFTAIAPTRAVLDRWRGVNKQTAMKHLLSFKKELGTLTSAINRRSSKQ
                     KKRGGKTGIAVMIGLIASVGAVTLSNFQGKVMMTVNATDVTDVITIPTAAGKNLCIVR
                     AMDVGYMCDDTITYECPVLSAGNDPEDIDCWCTKSAVYVRYGRCTKTRHSRRSRRSLT
                     VQTHGESTLANKKGAWMDSTKATRYLVKTESWILRNPGYALVAAVIGWMLGSNTMQRV
                     VFVVLLLLVAPAYSFNCLGMSNRDFLEGVSGATWVDLVLEGDSCVTIMSKDKPTIDVK
                     MMNMEAANLAEVRSYCYLATVSDLSTKAACPTMGEAHNDKRADPAFVCRQGVVDRGWG
                     NGCGLFGKGSIDTCAKFACSTKAIGRTILKENIKYEVAIFVHGPTTVESHGNYSTQVG
                     ATQAGRFSITPAAPSHTLKLGEYGEVTVDCEPRSGIDTNAYVMTVGTKTFLVHREWF
                     MDLNLPWSSAGSTVWRNRETLMEFEEPHATKQSVIALGSQEGALHQALAGAIPVEFSS
```

NTVKLTSGHLKCRVKMEKLQLKGTTYGVCSKAFKFLGTPADTGHGTVVLELQYTGTDG
PCKVPISSVASLNDLTPVGRLVTVNPFVSVATANAKVLIELEPPFGDSYIVVGRGEQQ
INHHWHKSGSSIGKAFTTTLKGAQRLAALGDTAWDFGSVGGVFTSVGKAVHQVFGGAF
RSLFGGMSWITQGLLGALLLWMGINARDRSIALTFLAVGGVLLFLSVNVHADTGCAID
ISRQELRCGSGVFIHNDVEAWMDRYKYYPETPQGLAKIIQKAHKEGVCGLRSVSRLEH
QMWEAVKDELNTLLKENGVDLSVVVEKQEGMYKSAPKRLTATTEKLEIGWKAWGKSIL
FAPELANNTFVVDGPETKECPTQNRAWNSLEVEDFGFGLTSTRMFLKVRESNTTECDS
KIIGTAVKNNLAIHSDLSYWIESRLNDTWKLERAVLGEVKSCTWPETHTLWGDGILES
DLIIPVTLAGPRSNHNRRPGYKTQNQGPWDEGRVEIDFDYCPGTTVTLSESCGHRGPA
TRTTTESGKLITDWCCRSCTLPPLRYQTDSGCWYGMEIRPQRHDEKTLVQSQVNAYNA
DMIDPFQLGLLVVFLATQEVLRKRWTAKISLPAILIALLVLVFGGITYTDVLRYVILV
GAAFAESNSGGDVVHLALMATFKIQPVFMVASFLKARWTNQENILLMLAAVFFQMAYH
DARQILLWEIPDVLNSLAVAWMILRAITFTTTSNVVVPLLALLTPGLRCLNLDVYRIL
LLMVGIGSLIREKRSAAAKKKGASLLCLALASTGLFNPMILAAGLIACDPNRKRGWPA
TEVMTAVGLMFAIVGGLAELDIDSMAIPMTIAGLMFAAFVISGKSTDMWIERTADISW
ESDAEITGSSERVDVRLDDDGNFQLMNDPGAPWKIWMLRMVCLAISAYTPWAILPSVV
GFWITLQYTKRGGVLWDTPSPKEYKKGDTTTGVYRIMTRGLLGSYQAGAGVMVEGVFH
TLWHTTKGAALMSGEGRLDPYWGSVKEDRLCYGGPWKLQHKWNGQDEVQMIVVEPGKN
VKNVQTKPGVFKTPEGEIGAVTLDFPTGTSGSPIVDKNGDVIGLYGNGVIMPNGSYIS
AIVQGERMDEPIPAGFEPEMLRKKQITVLDLHPGAGKTRRILPQIIKEAINRRLRTAV
LAPTRVVAAEMAEALRGLPIRYQTSAVPREHNGNEIVDVMCHATLTHRLMSPHRVPNY
NLFVMDEAHFTDPASIAARGYISTKVELGEAAAIFMTATPPGTSDPFPESNSPISDLQ
TEIPDRAWNSGYEWITEYTGKTVWFVPSVKMGNEIALCLQRAGKKVVQLNRKSYETEY
PKCKNDDWDFVITTDISEMGANFKASRVIDSRKSVKPTIITEGEGRVILGEPSAVTAA
SAAQRRGRIGRNPSQVGDEYCYGGHTNEDDSNFAHWTEARIMLDNINMPNGLIAQFYQ
PEREKVYTMDGEYRLRGEERKNFLELLRTADLPVWLAYKVAAAGVSYHDRRWCFDGPR
TNTILEDNNEVEVITKLGERKILRPRWIDARVYSDHQALKAFKDFASGKRSQIGLIEV
LGKMPEHFMGKTWEALDTMYVVATAEKGGRAHRMALEELPDALQTIALIALLSVMTMG
VFFLLMQRKGIGKIGLGGAVLGVATFFCWMAEVPGTKIAGMLLLSLLLMIVLIPEPEK
QRSQTDNQLAVFLICVMTLVSAVAANEMGWLDKTKSDISSLFGQRIEVKENFSMGEFL
LDLRPATAWSLYAVTTAVLTPLLKHLITSDYINTSLTSINVQASALFTLARGFPFVDV
GVSALLLAAGCWGQVTLTVTVTAATLLFCHYAYMVPGWQAEAMRSAQRRTAAGIMKNA
VVDGIVATDVPELERTTPIMQKKVGQIMLILVSLAAVVVNPSVKTVREAGILITAAAV
TLWENGASSVWNATTAIGLCHIMRGGWLSCLSITWTLIKNMEKPGLKRGGAKGRTLGE
VWKERLNQMTKEEFTRYRKEAIIEVDRSAAKHARKEGNVTGGHPVSRGTAKLRWLVER
RFLEPVGKVIDLGCGRGGWCYYMATQKRVQEVRGYTKGGPGHEEPQLVQSYGWNIVTM
KSGVDVFYRPSECCDTLLCDIGESSSSAEVEEHRTIRVLEMVEDWLHRGPREFCVKVL
CPYMPKVIEKMELLQRRYGGGLVRNPLSRNSTHEMYWVSRASGNVVHSVNMTSQVLLG
RMEKRTWKGPQYEEDVNLGSGTRAVGKPLLNSDTSKIKNRIERLRREYSSTWHHDENH
PYRTWNYHGSYDVKPTGSASSLVNGVVRLLSKPWDTITNVTTMAMTDTTPFGQQRVFK
EKVDTKAPEPPEGVKYVLNETTNWLWAFLAREKRPRMCSREEFIRKVNSNAALGAMFE
EQNQWRSAREAVEDPKFWEMVDEEREAHLRGECHTCIYNMMGKREKKPGEFGKAKGSR
AIWFMWLGARFLEFEALGFLNEDHWLGRKNSGGGVEGLGLQKLGYILREVGTRPGGKI
YADDTAGWDTRITRADLENEAKVLELLDGEHRRLARAIIELTYRHKVVKVMRPAADGR
TVMDVISREDQRGSGQVVTYALNTFTNLAVQLVRMMEGEGVIGPDDVEKLTKGKGPKV
RTWLFENGEERLSRMAVSGDDCVVKPLDDRFATSLHFLNAMSKVRKDIQEWKPSTGWY
DWQQVPFCSNHFTELIMKDGRTLVVPCRGQDELVGRARISPGAGWNVRDTACLAKSYA
QMWLLLYFHRRDLRLMANAICSAVPVNWVPTGRTTWSIHAGGEWMTTEDMLEVWNRVW
IEENEWMEDKTPVEKWSDVPYSGKREDIWCGSLIGTRARATWAENIQVAINQVRAIIG
DEKYVDYMSSLKRYEDTTLVEDTVL"
ORIGIN
        1 agtagttcgc ctgtgtgagc tgacaaactt agtagtgttt gtgaggatta acaacaatta
       61 acacagtgcg agctgtttct tagcacgaag atctcgatgt ctaagaaacc aggagggccc
      121 ggcaagagcc gggctgtcaa tatgctaaaa cgcggaatgc cccgcgtgtt gtccttgatt
      181 ggactgaaga gggctatgtt gagcctgatc gacggcaagg ggccaatacg atttgtgttg
      241 gctctcttgg cgttcttcag gttcacagca attgctccga cccgagcagt gctggatcga
      301 tggagaggtg tgaacaaaca aacagcgatg aaacaccttc tgagttttaa gaaggaacta
      361 gggaccttga ccagtgctat caatcggcgg agctcaaaac aaaagaaaag aggaggaaag
      421 accggaattg cagtcatgat tggcctgatc gccagcgtag gagcagttac cctctctaac
      481 ttccaaggga aggtgatgat gacggtaaat gctactgacg tcacagatgt catcacgatt
      541 ccaacagctg ctggaaagaa cctatgcatt gtcagagcaa tggatgtggg atacatgtgc
      601 gatgatacta tcacttatga atgcccagtg ctgtcggctg gtaatgatcc agaagacatc
      661 gactgttggt gcacaaagtc agcagtctac gtcaggtatg gaagatgcac caagacacgc
      721 cactcaagac gcagtcggag gtcactgaca gtgcagacac acggagaaag cactctagcg
      781 aacaagaagg gggcttggat ggacagcacc aaggccacaa ggtatttggt aaaaacagaa
      841 tcatggatct tgaggaaccc tggatatgcc ctggtggcag ccgtcattgg ttggatgctt

```
 901 gggagcaaca ccatgcagag agttgtgttt gtcgtgctat tgcttttggt ggccccagct
 961 tacagcttca actgccttgg aatgagcaac agagacttct tggaaggagt gtctggagca
1021 acatgggtgg atttggttct cgaaggcgac agctgcgtga ctatcatgtc taaggacaag
1081 cctaccatcg atgtgaagat gatgaatatg gaggcggcca acctggcaga ggtccgcagt
1141 tattgctatt tggctaccgt cagcgatctc tccaccaaag ctgcgtgccc gaccatggga
1201 gaagctcaca atgacaaacg tgctgaccca gcttttgtgt cagacaagg agtggtggac
1261 aggggctggg gcaacggctg cggactattt ggcaaaggaa gcattgacac atgcgccaaa
1321 tttgcctgct ctaccaaggc aataggaaga accatcttga aagagaatat caagtacgaa
1381 gtggccattt ttgtccatgg accaactact gtggagtcgc acggaaacta ctccacacag
1441 gttggagcca ctcaggcagg gagattcagc atcactcctg cggcgccttc acacacacta
1501 aagcttggag aatatggaga ggtgacagtg gactgtgaac cacggtcagg gattgacacc
1561 aatgcatact acgtgatgac tgttggaaca aagacgttct tggtccatcg tgagtggttc
1621 atggacctca acctcccttg gagcagtgct ggaagtactg tgtggaggaa cagagacg
1681 ttaatggagt ttgaggaacc acacgccacg aagcagtctg tgatagcatt gggctcacaa
1741 gagggagctc tgcatcaagc tttggctgga gccattcctg tggaatttc aagcaacact
1801 gtcaagttga cgtcgggtca tttgaagtgt agagtgaaga tggaaaaatt gcagttgaag
1861 ggaacaacct atggcgtctg ttcaaaggct ttcaagtttc ttgggactcc cgcagacaca
1921 ggtcacggca ctgtggtgtt ggaattgcag tacactggca cggatggacc ttgcaaagtt
1981 cctatctcgt cagtggcttc attgaacgac ctaacgccag tgggcagatt ggtcactgtc
2041 aacccttttg tttcagtggc cacggccaac gctaaggtcc tgattgaatt ggaaccaccc
2101 tttggagact catacatagt ggtgggcaga ggagaacaac agatcaatca ccattggcac
2161 aagtctggaa gcagcattgg caaagccttt acaaccaccc tcaaaggagc gcagagacta
2221 gccgctctag gagacacagc ttgggacttt ggatcagttg gaggggtgtt cacctcagtt
2281 gggaaggctg tccatcaagt gttcggagga gcattccgct cactgttcgg aggcatgtcc
2341 tggataacgc aaggattgct gggggctctc ctgttgtgga tgggcatcaa tgctcgtgat
2401 aggtccatag ctctcacgtt tctcgcagtt ggaggagttc tgctcttcct ctccgtgaac
2461 gtgcacgctg acactgggtg tgccatagac atcacgcggc aagagctgag atgtggaagt
2521 ggagtgttca cacacaatga tgtggaggct tggatggacc ggtacaagta ttaccctgaa
2581 acgccacaag gcctagccaa gatcattcag aaagctcata aggaaggagt gtgcggtcta
2641 cgatcagttt ccagactgga gcatcaaatg tgggaagcag tgaaggacga gctgaacact
2701 cttttgaagg agaatggtgt ggaccttagt gtcgtggttg agaaacagga gggaatgtac
2761 aagtcagcac ctaaacgcct caccgccacc acggaaaaat tggaaattgg ctggaaggcc
2821 tggggaaaga gtatttttatt tgcaccagaa ctcgccaaca cacctttgt ggttgatggt
2881 ccggagacca aggaatgtcc gactcagaat cgcgcttgga atagcttaga agtggaggat
2941 tttggattg gtctcaccag cactcggatg ttcctgaagg tcagagagag caacacaact
3001 gaatgtgact cgaagatcat tggaacggct gtcaagaaca acttggcgat ccacagtgac
3061 ctgtcctatt ggattgaaag caggctcaat gatacgtgga agcttgaaag ggcagttctg
3121 ggtgaagtca aatcatgtac gtggcctgag acgcatacct tgtggggcga tggaatcctt
3181 gagagtgact tgataatacc agtcacactg gcgggaccac gaagcaatca caatcggaga
3241 cctgggtaca agacacaaaa ccagggcccg tgggacgaag gccgggtaga gattgacttc
3301 gattactgcc caggaaccac ccacagagag cggaaagttg ataacgatt ggtgctgcag
3361 actcgcacca ccacagagag cggaaagttg ataacgatt ggtgctgcag gagctgcacc
3421 ttaccaccac tgcgctacca aactgacagc ggctgttggt atggtatgga gatcagacca
3481 cagagacatg atgaaaagac cctcgtgcag tcacaagtga atgcttataa tgctgatatg
3541 attgacccct ttcagttggg ccttctggtc gtgttcttgg ccacccagga ggtccttcgc
3601 aagaggtgga cagccaagat cagcttgcca gctatactga ttgctctgct agtcctggtg
3661 tttgggggca ttacttacac tgatgtgtta cgctatgtca tcttggtggg ggcagctttc
3721 gcagaatcta attcggagg agacgtggta cacttggcgc tcatggcgac ctttaagata
3781 caaccagtgt ttatggtggc atcgtttctc aaagcgagat ggaccaacca ggagaacatt
3841 ttgttgatgt tggcggctgt tttctttcaa atggcttatc acgatgcccg ccaaattctg
3901 ctctgggaga tccctgatgt gttgaattca ctggcggtag cttggatgat actgagagcc
3961 ataacattca caacgacatc aaacgtggtt gttccgctgc tagccctgct aacacccggg
4021 ctgagatgct tgaatctgga tgtgtacagg atactgctgt gatggtcgg aataggcagc
4081 ttgatcaggg agagaggag tgcagctgca aaaaagaaag gagcaagtct gctatgcttg
4141 gctctagcct caacaggact tttcaacccc atgatccttg ctgctggact gattgcatgt
4201 gatcccaacc gtaaacgcgg atggcccgca actgaagtga tgacagctgt cggcctaatg
4261 tttgccatcg tcggagggct ggcagagctt gacattgact ccatggccat tccaatgact
4321 atcgcggggc tcatgtttgc tgctttcgtg atttctggga aatcaacaga tatgtggatt
4381 gagagaacgg cggacatttc ctgggaaagt gatgcagaaa ttacaggctc gagcgaaaga
4441 gttgatgtgc ggcttgatga tgatggaaac ttccagctca tgaatgatcc aggagcacct
4501 tggaagatat ggatgctcag aatggtctgt ctcgcgatta gtgcgtacac cccctgggca
4561 atcttgccct cagtagttgg attttggata actctccaat acacaaagag aggaggcgtg
4621 ttgtgggaca ctccctcacc aaaggagtac aaaaagggg acacgaccac cggcgtctac
4681 aggatcatga ctcgtgggct gctcggcagt tatcaagcag agcgggcgt gatggttgaa
4741 ggtgtttttcc acaccctttg gcatacaaca aaaggagccg ctttgatgag cggagagggc
4801 cgcctggacc catactgggg cagtgtcaag gaggatcgac tttgttacgg aggaccctgg
```

```
4861 aaattgcagc acaagtggaa cgggcaggat gaggtgcaga tgattgtggt ggaacctggc
4921 aagaacgtta agaacgtcca gacgaaacca ggggtgttca aaacacctga aggagaaatc
4981 ggggccgtga ctttggactt ccccactgga acatcaggct caccaatagt ggacaaaaac
5041 ggtgatgtga ttgggcttta tggcaatgga gtcataatgc ccaacggctc atacataagc
5101 gcgatagtgc agggtgaaag gatggatgag ccaatcccag ccggattcga acctgagatg
5161 ctgaggaaaa aacagatcac tgtactggat ctccatcccg gcgccggtaa aacaaggagg
5221 attctgccac agatcatcaa agaggccata aacagaagac tgagaacagc cgtgctagca
5281 ccaaccaggg ttgtggctgc tgagatggct gaagcactga gaggactgcc catccggtac
5341 cagacatccg cagtgcccag agaacataat ggaaatgaga ttgttgatgt catgtgtcat
5401 gctaccctca cccacaggct gatgtctcct cacagggtgc cgaactacaa cctgttcgtg
5461 atggatgagg ctcatttcac cgacccagct agcattgcag caagaggtta catttccaca
5521 aaggtcgagc taggggaggc ggcggcaata ttcatgacag ccacccacc aggcacttca
5581 gatccattcc cagagtccaa ttcaccaatt tccgacttac agactgagat cccggatcga
5641 gcttggaact ctggatacga atggatcaca gaatacaccg ggaagacggt ttggtttgtg
5701 cctagtgtta agatggggaa tgagattgcc ctttgcctac aacgtgctgg aaagaaagta
5761 gtccaattga acagaaagtc gtacgagacg gagtacccaa aatgtaagaa cgatgattgg
5821 gactttgtta tcacaacaga catatctgaa atgggggcta acttcaaggc gagcagggtg
5881 attgacagcc ggaagagtgt gaaaccaacc atcataacag aaggagaagg gagagtgatc
5941 ctgggagaac catctgcagt gacagcagct agtgccgccc agagacgtgg acgtatcggt
6001 agaaatccgt cgcaagttgg tgatgagtac tgttatgggg ggcacacgaa tgaagacgac
6061 tcgaacttcg cccattggac tgaggcacga atcatgctgg acaacatcaa catgccaaac
6121 ggactgatcg ctcaattcta ccaaccagag cgtgagaagg tatataccat ggatggggaa
6181 taccggctca gaggagaaga gagaaaaaac tttctggaac tgttgaggac tgcagatctg
6241 ccggtttggc tggcttacaa ggttgcagcg gctggagtgt cataccacga ccggaggtgg
6301 tgctttgatg gtcctaggac aaacacaatt ttagaagaca caacgaagt ggaagtcatc
6361 acgaagcttg gtgaaaggaa gattctgagg ccgcgctgga ttgacgccag ggtgtactcg
6421 gatcaccagg cactaaaggc gttcaaggac ttcgcctcgg gaaaacgttc tcagataggg
6481 ctcattgagg ttctgggaaa gatgcctgag cacttcatgg ggaagacatg ggaagcactt
6541 gacaccatgt acgttgtggc cactgcagag aaaggaggaa gagctcacag aatggccctg
6601 gaggaactgc cagatgctct tcagacaatt gccttgattg ccttattgag tgtgatgacc
6661 atgggagtat tcttcctcct catgcagcgg aagggcattg gaaagatagg tttgggaggc
6721 gctgtcttgg gagtcgcgac ctttttctgt tggatggctg aagttccagg aacgaagatc
6781 gccggaatgt tgctgctctc ccttctcttg atgattgtgc taattcctga gccagagaag
6841 caacgttcgc agacagacaa ccagctagcc gtgttcctga tttgtgtcat gacccttgtg
6901 agcgcagtgg cagccaacga gatgggttgg ctagataaga ccaagagtga cataagcagt
6961 ttgtttgggc aaagaattga ggtcaaggag aatttcagca tgggagagtt tcttctggac
7021 ttgaggccgg caacagcctg gtcactgtac gctgtgacaa cagcggtcct cactccactg
7081 ctaaagcatt tgatcacgtc agattacatc aacacctcat tgacctcaat aaacgttcag
7141 gcaagtgcac tattcacact cgcgcgaggc ttccccttcg tcgatgttgg agtgtcggct
7201 ctcctgctag cagccggatg ctggggacaa gtcaccctca ccgttacggt aacagcggca
7261 acactccttt tttgccacta tgcctacatg gttcccggtt ggcaagctga ggcaatgcgc
7321 tcagcccagc ggcggacagc ggccggaatc atgaagaacg ctgtagtgga tggcatcgtg
7381 gccacggacg tcccagaatt agagcgcacc acacccatca tgcagaagaa agttggacag
7441 atcatgctga tcttggtgtc tctagctgca gtagtagtga accgtctgt gaagacagta
7501 cgagaagccg gaattttgat cacggccgca gcggtgacgc tttgggagaa tggagcaagc
7561 tctgtttgga acgcaacaac tgccatcgga ctctgccaca tcatgcgtgg gggttggttg
7621 tcatgtctat ccataacatg gacactcata aagaacatgg aaaaaccagg actaaaaaga
7681 ggtggggcaa aaggacgcac cttgggagag gttttggaaag aaagactcaa ccagatgaca
7741 aaagaagagt tcactaggta ccgcaaagag gccatcatcg aagtcgatcg ctcagcggca
7801 aaacacgcca ggaaagaagg caatgtcact ggagggcatc cagtctctag gggcacagca
7861 aaactgagat ggctggtcga acggaggttt ctcgaaccgg tcggaaaagt gattgacctt
7921 ggatgtggaa gaggcggttg gtgttactat atggcaaccc aaaaaagagt ccaagaagtc
7981 agagggtaca caaaggcgg tcccggacat gaagagcccc aactagtgca aagttatgga
8041 tggaacattg tcaccatgaa gagtggagtg gatgtgttct acagaccttc tgagtgttgt
8101 gacaccctcc tttgtgacat cggagagtcc tcgtcaagtg ctgaggttga agacagatagg
8161 acgattcggg tccttgaaat ggttgaggac tggctgcacc gagggccaag ggaattttgc
8221 gtgaaggtgc tctgcccca catgccgaaa gtcatagaga agatggagct gctccaacgc
8281 cggtatgggg ggggactggt cagaaaccca ctctcacgga attccacgca cgagatgtat
8341 tgggtgagtc gagcttcagg caatgtggta cattcagtga atatgaccag ccaggtgctc
8401 ctaggaagaa tggaaaaaag gacctggaag ggaccccaat acgaggaaga tgtaaacttg
8461 ggaagtggaa ccagggcggt gggaaaaccc ctgctcaact cagacaccag taaaatcaag
8521 aacaggattg aacgactcag gcgtgagtac agttcgacgt ggcaccacga tgagaaccac
8581 ccatatagaa cctggaacta tcacggcagt tatgatgtga gcccacagg ctccgccagt
8641 tcgctggtca atggagtggt caggctcctc tcaaaaccat gggacaccat cacgaatgtt
8701 accaccatgg ccatgactga cactactccc ttcgggcagc agcgagtgtt caaagagaag
8761 gtggacacga aagctcctga accgccagaa ggagtgaagt atgtgctcaa tgagaccacc
```

```
 8821 aactggttgt gggcgttttt ggccagagaa aaacgtccca gaatgtgctc tcgagaggaa
 8881 ttcataagaa aggtcaacag caatgcagct ttgggtgcca tgtttgaaga gcagaatcaa
 8941 tggaggagcg ccagagaagc agttgaagat ccaaaatttt gggagatggt ggatgaggag
 9001 cgcgaggcac atctgcgggg ggaatgtcac acttgcattt acaacatgat gggaaagaga
 9061 gagaaaaaac ccggagagtt cggaaaggcc aagggaagca gagccatttg gttcatgtgg
 9121 ctcggagctc gctttctgga gttcgaggct ctgggttttc tcaatgaaga ccactggctt
 9181 ggaagaaaga actcaggagg aggtgtcgag ggcttgggcc tccaaaaact gggttacatc
 9241 ctgcgtgaag ttggcacccg gcctgggggc aagatctatg ctgatgacac agctggctgg
 9301 gacacccgca tcacgagagc tgacttggaa aatgaagcta aggtgcttga gctgcttgat
 9361 ggggaacatc ggcgtcttgc cagggccatc attgagctca cctatcgtca caaagttgtg
 9421 aaagtgatgc gcccggctgc tgatggaaga accgtcatgg atgttatctc cagagaagat
 9481 cagaggggga gtggacaagt tgtcacctac gccctaaaca ctttcaccaa cctggccgtc
 9541 cagctggtga ggatgatgga aggggaagga gtgattggcc cagatgatgt ggagaaactc
 9601 acaaaaggga aaggacccaa agtcaggacc tggctgtttg agaatgggga agaaagactc
 9661 agccgcatgg ctgtcagtgg agatgactgt gtggtaaagc ccctggacga tcgctttgcc
 9721 acctcgctcc acttcctcaa tgctatgtca aaggttcgca aagacatcca agagtggaaa
 9781 ccgtcaactg gatggtatga ttggcagcag gttccatttt gctcaaacca tttcactgaa
 9841 ttgatcatga aagatggaag aacactggtg gttccatgcc gaggacagga tgaattggta
 9901 ggcagagctc gcatatctcc aggggccgga tggaacgtcc gcgacactgc ttgtctggct
 9961 aagtcttatg cccagatgtg gctgcttctg tacttccaca gaagagacct gcggctcatg
10021 gccaacgcca tttgctccgc tgtccctgtg aattgggtcc ctaccggaag aaccacgtgg
10081 tccatccatg caggaggaga gtggatgaca acagaggaca tgttggaggt ctggaaccgt
10141 gtttggatag aggagaatga atggatggaa gacaaaaccc cagtggagaa atggagtgac
10201 gtcccatatt caggaaaacg agaggacatc tggtgtggca gcctgattgg cacaagagcc
10261 cgagccacgt gggcagaaaa catccaggtg gctatcaacc aagtcagagc aatcatcgga
10321 gatgagaagt atgtggatta tatgagttca ctaaagagat atgaagacac aactttggtt
10381 gaggacacag tactgtagat atttaattaa ttgtaaatag acaatataag tatgcataaa
10441 agtgtagttt tatagtagta tttagtggtg ttagtgtaaa tagttaagaa aattttgagg
10501 agaaagtcag gccgggaagt tcccgccacc ggaagttgag tagacggtgc tgcctgcgac
10561 tcaaccccag gaggactggg tgaacaaagc cgcgaagtga tccatgtaag ccctcagaac
10621 cgtctcggaa ggaggacccc acatgttgta acttcaaagc ccaatgtcag accacgctac
10681 ggcgtgctac tctgcggaga gtgcagtctg cgatagtgcc ccaggaggac tgggttaaca
10741 aaggcaaacc aacgccccac gcggccctag ccccggtaat ggtgttaacc agggcgaaag
10801 gactagaggt tagaggagac cccgcggttt aaagtgcacg gcccagcctg gctgaagctg
10861 taggtcaggg gaaggactag aggttagtgg agaccccgtg ccacaaaaca ccacaacaaa
10921 acagcatatt gacacctggg atagactagg agatcttctg ctctgcacaa ccagccacac
10981 ggcacagtgc gccgacaatg gtggctggtg gtgcgagaac acaggatct
```
//

Apr 17 2007 11:10:07

PubMed   Nucleotide   Protein   Genome   Structure   PMC   Taxonomy   OMIM   Books

My NCBI
[Sign In] [Register]

Search CoreNucleotide   for   Go   Clear

Limits   Preview/Index   History   Clipboard   Details

Display GenBank   Show 5   Send to   Hide: ☐ sequence   ☐ all but gene, CDS and mRNA features

Range: from begin   to end   ☐ Reverse complemented strand   Features: [+]   Refresh

☐ 1: AF404754. Reports West Nile virus i...[gi:21929234]   Links

Features   Sequence

```
LOCUS       AF404754                11029 bp ss-RNA     linear   VRL 23-JUL-2002
DEFINITION  West Nile virus isolate WN NJ 2000 MQ5488, complete genome.
ACCESSION   AF404754
VERSION     AF404754.1  GI:21929234
KEYWORDS    .
SOURCE      West Nile virus
  ORGANISM  West Nile virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
            Flavivirus; Japanese encephalitis virus group.
REFERENCE   1  (bases 1 to 11029)
  AUTHORS   Lanciotti,R.S., Ebel,G.D., Deubel,V., Kerst,A.J., Murri,S.,
            Meyer,R., Bowen,M., McKinney,N., Morrill,W.E., Crabtree,M.B.,
            Kramer,L.D. and Roehrig,J.T.
  TITLE     Complete genome sequences and phylogenetic analysis of West Nile
            virus strains isolated from the United States, Europe, and the
            Middle East
  JOURNAL   Virology 298 (1), 96-105 (2002)
   PUBMED   12093177
REFERENCE   2  (bases 1 to 11029)
  AUTHORS   Lanciotti,R.S., Ebel,G.D. and Kerst,A.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Division of Vector-Borne Infectious
            Diseases, Centers for Disease Control & Prevention, Rampart Road,
            Fort Collins, CO 80521, USA
FEATURES             Location/Qualifiers
     source          1..11029
                     /organism="West Nile virus"
                     /mol_type="genomic RNA"
                     /isolate="WN NJ 2000 MQ5488"
                     /specific_host="Culex pipiens pool"
                     /db_xref="taxon:11082"
                     /country="USA: New Jersey"
     CDS             97..10398
                     /note="contains capsid, pre-membrane, envelope, NS1, NS2a,
                     NS2b, NS3, NS4a, NS4b, and NS5"
                     /codon_start=1
                     /product="polyprotein precursor"
                     /protein_id="AAM81750.1"
                     /db_xref="GI:21929235"
                     /translation="MSKKPGGPGKSRAVNMLKRGMPRVLSLIGLKRAMLSLIDGKGPI
                     RFVLALLAFFRFTAIAPTRAVLDRWRGVNKQTAMKHLLSFKKELGTLTSAINRRSSKE
                     KKRGGKTGIAVMIGLIASVGAVTLSNFQGKVMMTVNATDVTDVITIPTAAGKNLCIVR
                     AMDVGYMCDDTITYECPVLSAGNDPEDIDCWCTKSAVYVRYGRCTKTRHSRRSRRSLT
                     VQTHGESTLANKKGAWMDSTKATRYLVKTESWILRNPGYALVAAVIGWMLGSNTMQRV
                     VFVVLLLLVAPAYSFNCLGMSNRDFLEGVSGATWVDLVLEGDSCVTIMSKDKPTIDVK
                     MMNMEAANLAEVRSYCYLATVSDLSTKAACPTMGEAHNDKRADPAFVCRQGVVDRGWG
                     NGCGLFGKGSIDTCAKFACSTKAIGRTILKENIKYEVAIFVHGPTTVESHGNYSTQVG
                     ATQAGRFSITPAAPSYTLKLGEYGEVTVDCEPRSGIDTNAYYVMTVGTKTFLVHREWF
                     MDLNLPWSSAGSTVWRNRETLMEFEEPHATKQSVIALGSQEGALHQALAGAIPVEFSS
```

```
                    NTVKLTSGHLKCRVKMEKLQLKGTTYGVCSKAFKFLGTPADTGHGTVVLELQYTGTDG
                    PCKVPISSVASLNDLTPVGRLVTVNPFVSVATANAKVLIELEPPFGDSYIVVGRGEQQ
                    INHHWHKSGSSIGKAFTTTLKGAQRLAALGDTAWDFGSVGGVFTSVGKAVHQVFGGAF
                    RSLFGGMSWITQGLLGALLLWMGINARDRSIALTFLAVGGVLLFLSVNVHADTGCAID
                    ISRQELRCGSGVFIHNDVEAWMDRYKYYPETPQGLAKIIQKAHKEGVCGLRSVSRLEH
                    QMWEAVKDELNTLLKENGVDLSVVVEKQEGMYKSAPKRLTATTEKLEIGWKAWGKSIL
                    FAPELANNTFVVDGPETKECPTQNRAWNSLEVEDFGFGLTSTRMFLKVRESNTTECDS
                    KIIGTAVKNNLAIHSDLSYWIESRLNDTWKLERAVLGEVKSCTWPETHTLWGDGILES
                    DLIIPVTLAGPRSNHNRRPGYKTQNQGPWDEGRVEIDFDYCPGTTVTLSESCGHRGPA
                    TRTTTESGKLITDWCCRSCTLPPLRYQTDSGCWYGMEIRPQRHDEKTLVQSQVNAYNA
                    DMIDPFQLGLLVVFLATQEVLRKRWTAKISMPAILIALLVLVFGGITYTDVLRYVILV
                    GAAFAESNSGGDVVHLALMATFKIQPVFMVASFLKARWTNQENILLMLAAVFFQMAYH
                    DARQILLWEIPDVLNSLAVAWMILRAITFTTTSNVVVPLLALLTPGLRCLNLDVYRIL
                    LLMVGIGSLIREKRSAAAKKKGASLLCLALASTGLFNPMILAAGLIACDPNRKRGWPA
                    TEVMTAVGLMFAIVGGLAELDIDSMAIPMTIAGLMFAAFVISGKSTDMWIERTADISW
                    ESDAEITGSSERVDVRLDDDGNFQLMNDPGAPWKIWMLRMVCLAISAYTPWAILPSVV
                    GFWITLQYTKRGGVLWDTPSPKEYKKGDTTTGVYRIMTRGLLGSYQAGAGVMVEGVFH
                    TLWHTTKGAALMSGEGRLDPYWGSVKEDRLCYGGPWKLQHKWNGQDEVQMIVVEPGKN
                    VKNVQTKPGVFKTPEGEIGAVTLDFPTGTSGSPIVDKNGDVIGLYGNGVIMPNGSYIS
                    AIVQGERMDEPIPAGFEPEMLRKKQITVLDLHPGAGKTRRILPQIIKEAINRRLRTAV
                    LAPTRVVAAEMAEALRGLPIRYQTSAVPREHNGNEIVDVMCHATLTHRLMSPHRVPNY
                    NLFVMDEAHFTDPASIAARGYISTKVELGEAAAIFMTATPPGTSDPFPESNSPISDLQ
                    TEIPDRAWNSGYEWITEYTGKTVWFVPSVKMGNEIALCLQRAGKKVVQLNRKSYETEY
                    PKCKNDDWDFVITTDISEMGANFKASRVIDSRKSVKPTIITEGEGRVILGEPSAVTAA
                    SAAQRRGRIGRNPSQVGDEYCYGGHTNEDDSNFAHWTEARIMLDNINMPNGLIAQFYQ
                    PEREKVYTMDGEYRLRGEERKNFLELLRTADLPVWLAYKVAAAGVSYHDRRWCFDGPR
                    TNTILEDNNEVEVITKLGERKILRPRWIDARVYSDHQALKAFKDFASGKRSQIGLIEV
                    LGKMPEHFMGKTWEALDTMYVVATAEKGGRAHRMALEELPDALQTIALIALLSVMTMG
                    VFFLLMQRKGIGKIGLGGAVLGVATFFCWMAEVPGTKIAGMLLLSLLLMIVLIPEPEK
                    QRSQTDNQLAVFLICVMTLVSAVAANEMGWLDKTKSDISSLFGQRIEVKENFSMGEFL
                    LDLRPATAWSLYAVTTAVLTPLLKHLITSDYINTSLTSINVQASALFTLARGFPFVDV
                    GVSALLLAAGCWGQVTLTVTVTAATLLFCHYAYMVPGWQAEAMRSAQRRTAAGIMKNA
                    VVDGIVATDVPELERTTPIMQKKVGQIMLILVSLAAVVVNPSVKTVREAGILITAAAV
                    TLWENGASSVWNATTAIGLCHIMRGGWLSCLSITWTLIKNMEKPGLKRGGAKGRTLGE
                    VWKERLNQMTKEEFTRYRKEAIIEVDRSAAKHARKEGNVTGGHPVSRGTAKLRWLVER
                    RFLEPVGKVIDLGCGRGGWCYYMATQKRVQEVRGYTKGGPGHEEPQLVQSYGWNIVTM
                    KSGVDVFYRPSECCDTLLCDIGESSSSAEVEEHRTIRVLEMVEDWLHRGPREFCVKVL
                    CPYMPKVIEKMELLQRRYGGGLVRNPLSRNSTHEMYWVSRASGNVVHSVNMTSQVLLG
                    RMEKRTWKGPQYEEDVNLGSGTRAVGKPLLNSDTSKIKNRIERLRREYSSTWHHDENH
                    PYRTWNYHGSYDVKPTGSASSLVNGVVRLLSKPWDTITNVTTMAMTDTTPFGQQRVFK
                    EKVDTKAPEPPEGVKYVLNETTNWLWAFLAREKRPRMCSREEFIRKVNSNAALGAMFE
                    EQNQWRSAREAVEDPKFWEMVDEEREAHLRGECHTCIYNMMGKREKKPGEFGKAKGSR
                    AIWFMWLGARFLEFEALGFLNEDHWLGRKNSGGGVEGLGLQKLGYILREVGTRPGGKI
                    YADDTAGWDTRITRADLENEAKVLELLDGEHRRLARAIIELTYRHKVVKVMRPAADGR
                    TVMDVISREDQRGSGQVVTYALNTFTNLAVQLVRMMEGEGVIGPDDVEKLTKGKGPKV
                    RTWLFENGEERLSRMAVSGDDCVVKPLDDRFATSLHFLNAMSKVRKDIQEWKPSTGWY
                    DWQQVPFCSNHFTELIMKDGRTLVVPCRGQDELVGRARISPGAGWNVRDTACLAKSYA
                    QMWLLLYFHRRDLRLMANAICSAVPVNWVPTGRTTWSIHAGGEWMTTEDMLEVWNRVW
                    IEENEWMEDKTPVEKWSDVPYSGKREDIWCGSLIGTRARATWAENIQVAINQVRAIIG
                    DEKYVDYMSSLKRYEDTTLVEDTVL"
ORIGIN
        1 agtagttcgc ctgtgtgagc tgacaaactt agtagtgttt gtgaggatta acaacaatta
       61 acacagtgcg agctgtttct tagcacgaag atctcgatgt ctaagaaacc aggagggccc
      121 ggcaagagcc gggctgtcaa tatgctaaaa cgcggaatgc cccgcgtgtt gtccttgatt
      181 ggactgaaga gggctatgtt gagcctgatc gacggcaagg ggccaatacg atttgtgttg
      241 gctctcttgg cgttcttcag gttcacagca attgctccga cccgagcagt gctggatcga
      301 tggagaggtg tgaacaaaca aacagcgatg aaacaccttc tgagttttaa gaaggaacta
      361 gggaccttga ccagtgctat caatcggcgg agctcaaaag aaaagaaaag aggaggaaag
      421 accggaattg cagtcatgat tggcctgatc gccagcgtag gagcagttac cctctctaac
      481 ttccaaggga aggtgatgat gacggtaaat gctactgacg tcacagatgt catcacgatt
      541 ccaacagctg ctggaaagaa cctatgcatt gtcagagcaa tggatgtggg atacatgtgc
      601 gatgatacta tcacttatga atgcccagtg ctgtcggctg taatgatcc agaagacatc
      661 gactgttggt gcacaaagtc agcagtctac gtcaggtatg aagatgcac caagacacgc
      721 cactcaagac gcagtcggag gtcactgaca gtgcagacac acggagaaag cactctagcg
      781 aacaagaagg gggcttggat ggacagcacc aaggccacaa ggtatttggt aaaaacagaa
      841 tcatggatct tgaggaaccc tggatatgcc ctggtggcag ccgtcattgg ttggatgctt
```

```
 901 gggagcaaca ccatgcagag agttgtgttt gtcgtgctat tgcttttggt ggccccagct
 961 tacagcttca actgccttgg aatgagcaac agagacttct tggaaggagt gtctggagca
1021 acatgggtgg atttggttct cgaaggcgac agctgcgtga ctatcatgtc taaggacaag
1081 cctaccatcg atgtgaagat gatgaatatg gaggcggcca acctggcaga ggtccgcagt
1141 tattgctatt tggctaccgt cagcgatctc tccaccaaag ctgcgtgccc gaccatggga
1201 gaagctcaca atgacaaacg tgctgaccca gcttttgtgt gcagacaagg agtggtggac
1261 aggggctggg gcaacggctg cggactattt ggcaaaggaa gcattgacac atgcgccaaa
1321 tttgcctgct ctaccaaggc aataggaaga accatcttga aagagaatat caagtacgaa
1381 gtggccattt ttgtccatgg accaactact gtggagtcgc acggaaacta ctccacacag
1441 gttggagcca ctcaggcagg gagattcagc atcactcctg cggcgccttc atacacacta
1501 aagcttggag aatatgtgag ggtgacagtg gactgtgaac cacggtcagg gattgacacc
1561 aatgcatact acgtgatgac tgttggaaca aagacgttct tggtccatcg tgagtggttc
1621 atggacctca acctcccttg gagcagtgct ggaagtactg tgtggaggaa cagagagacg
1681 ttaatggagt tggaggaacc acacgccacg aagcagtctg tgatagcatt gggctcacaa
1741 gagggagctc tgcatcaagc tttggctgga gccattcctg tggaattttc aagcaacact
1801 gtcaagttga cgtcgggtca tttgaagtgt agagtgaaga tggaaaaatt gcagttgaag
1861 ggaacaacct atggcgtctg ttcaaaggct ttcaagtttc ttgggactcc cgcagacaca
1921 ggtcacggca ctgtggtgtt ggaattgcag tacactggca cggatggacc ttgcaaagtt
1981 cctatctcgt cagtggcttc attgaacgac ctaacgccag tgggcagatt ggtcactgtc
2041 aacccttttg tttcagtggc cacggccaac gctaaggtcc tgattgaatt ggaaccaccc
2101 tttggagact catacatagt ggtgggcaga ggagaacaac agatcaatca tcattggcac
2161 aagtctggaa gcagcattgg caaagccttt acaaccaccc tcaaaggagc gcagagacta
2221 gccgctctag gagacacagc ttgggacttt ggatcagttg gaggggtgtt cacctcagtt
2281 gggaaggctg tccatcaagt gttcggagga gcattccgct cactgttcgg aggcatgtcc
2341 tggataacgc aaggattgct gggggctctc ctgttgtgga tgggcatcaa tgctcgtgat
2401 aggtccatag ctctcacgtt tctcgcagtt ggaggagttc tgctcttcct ctccgtgaac
2461 gtgcacgctg acactgggtg tgccatagac atcagccggc aagagctgag atgtggaagt
2521 ggagtgttca tacacaatga tgtggaggct tggatggacc ggtacaagta ttaccctgaa
2581 acgccacaag gcctagccaa gatcattcag aaagctcata aggaaggagt gtgcggtcta
2641 cgatcagttt ccagactgga gcatcaaatg tgggaagcag tgaaggacga gctgaacact
2701 cttttgaagg agaatggtgt ggaccttagt gtcgtggttg agaaacagga gggaatgtac
2761 aagtcagcac ctaaacgcct caccgccacc acgaaaaat tggaaattgg ctggaaggcc
2821 tggggaaaga gtattttatt tgcaccagaa ctcgccaaca cacctttgt ggttgatggt
2881 ccggagacca aggaatgtcc gactcagaat cgcgcttgga atagcttaga agtggaggat
2941 tttggatttg gtctcaccag cactcggatg ttcctgaagg tcagagagag caacacaact
3001 gaatgtgact cgaagatcat tggaacggct gtcaagaaca acttggcgat ccacagtgac
3061 ctgtcctatt ggattgaaag caggctcaat gatacgtgga agcttgaaag ggcagttctg
3121 ggtgaagtca aatcatgtac gtggcctgag acgcatacct tgtggggcga tggaatcctt
3181 gagagtgact tgataatacc agtcacactg gcgggaccac gaagcaatca caatcggaga
3241 cctgggtaca agacacaaaa ccagggccca tgggacgaag gccgggtaga gattgacttc
3301 gattactgcc caggaactac ggtcaccctg agtgagagct gcggacaccg tggacctgcc
3361 actcgcacca ccacagagag cggaaagttg ataacagatt ggtgctgcag gagctgcacc
3421 ttaccaccac tgcgctacca aactgacagc ggctgttggt atggtatgga gatcagacca
3481 cagagacatg atgaaaagac cctcgtgcag tcacaagtga atgcttataa tgctgatatg
3541 attgaccctt ttcagttggg ccttctggtc gtgttcttgg ccacccagga ggtccttcgc
3601 aagaggtgga cagccaagat cagcatgcca gctatactga ttgctctgct agtcctggtg
3661 tttgggggca ttacttacac tgatgtgtta cgctatgtca tcttggtggg ggcagctttc
3721 gcagaatcta attcgggagg agacgtggta cacttggcgc tcatggcgac cttcaagata
3781 caaccagtgt ttatggtggc atcgtttctc aaagcgagat ggaccaacca ggagaacatt
3841 ttgttgatgt tggcggctgt tttctttcaa atggcttatc acgatgcccg ccaaattctg
3901 ctctgggaga tccctgatgt gttgaattca ctggcggtag cttggatgat actgagagcc
3961 ataacattca caacgacatc aaacgtggtt gttccgctgc tagccctgct aacacccggg
4021 ctgagatgct tgaatctgga tgtgtacagg atactgctgt tgatggtcgg aataggcagc
4081 ttgatcaggg agaagaggag tgcagctgca aaaaagaaag gagcaagtct gctatgcttg
4141 gctctagcct caacaggact tttcaacccc atgatccttg ctgctggact gattgcatgt
4201 gatcccaacc gtaaacgcgg atggccccca actgaagtga tgacagctgt cggcctaatg
4261 tttgccatcg tcggagggct ggcagagctt gacattgact ccatggccat tccaatgact
4321 atcgcggggc tcatgtttgc tgctttcgtg atttctggga aatcaacaga catgtggatt
4381 gagagaacgg cggacatttc ctgggaaagt gatgcagaaa ttacaggctc gagcgaaaga
4441 gttgatgtgc ggcttgatga tgatggaaac ttccagctca tgaatgatcc aggagcacct
4501 tggaagatat ggatgctcag aatggtctgt ctcgcgatta gtgcgtacac ccctgggca
4561 atcttgccct cagtagttgg atttggata actctccaat acacaaagag aggaggcgtg
4621 ttgtgggaca ctccctcacc aaaggagtac aaaaaggggg acacgaccac cggcgtctac
4681 aggatcatga ctcgtgggct gctcggcagt tatcaagcag gagcgggcgt gatggttgaa
4741 ggtgttttcc acaccctttg gcatacaaca aaaggagccg ctttgatgag cggagagggc
4801 cgcctggacc catactgggg cagtgtcaag gaggatcgac tttgttacgg aggaccctgg
```

```
4861 aaattgcagc acaagtggaa cgggcaggat gaggtgcaga tgattgtggt ggaacctggc
4921 aagaacgtta agaacgtcca gacgaaacca ggggtgttca aaacacctga aggagaaatc
4981 ggggccgtga ctttggactt ccccactgga acatcaggct caccaatagt ggacaaaaac
5041 ggtgatgtga ttgggcttta tggcaatgga gtcataatgc ccaacggctc atacataagc
5101 gcgatagtgc agggtgaaag gatggatgag ccaatcccag ccggattcga acctgagatg
5161 ctgaggaaaa aacagatcac tgtactggat ctccatcccg gcgccggtaa aacaaggagg
5221 attctgccac agatcatcaa agaggccata aacagaagac tgagaacagc cgtgctagca
5281 ccaaccaggt tgtggctgc tgagatggct gaagcactga gaggactgcc catccggtac
5341 cagacatccg cagtgcccag agaacataat ggaaatgaga ttgttgatgt catgtgtcat
5401 gctaccctca cccacaggct gatgtctcct cacagggtgc cgaactacaa cctgttcgtg
5461 atggatgagg ctcatttcac cgacccagct agcattgcag caagaggtta catttccaca
5521 aaggtcgagc taggggaggc ggcggcaata ttcatgacag ccacccccacc aggcacttca
5581 gatccattcc cagagtccaa ttcaccaatt tccgacttac agactgagat cccggatcga
5641 gcttggaact ctggatacga atggatcaca gaatacaccg ggaagacggt ttggtttgtg
5701 cctagtgtca agatggggaa tgagattgcc ctttgcctac aacgtgctgga aaagaaagta
5761 gtccaattga acagaaagtc gtacgagacg gagtacccaa aatgtaagaa cgatgattgg
5821 gactttgtta tcacaacaga catatctgaa atgggggcta acttcaaggc gagcagggtg
5881 attgacagcc ggaagagtgt gaaaccaacc atcataacag aaggagaagg gagagtgatc
5941 ctgggagaac catctgcagt gacagcagct agtgccgccc agagacgtgg acgtatcggt
6001 agaaatccgt cgcaagttgg tgatgagtac tgttatgggg ggcacacgaa tgaagacgac
6061 tcgaacttcg cccattggac tgaggcacga atcatgctgg acaacatcaa catgccaaac
6121 ggactgatcg ctcaattcta ccaaccagag cgtgagaagg tatataccat ggatggggaa
6181 taccggctca gaggagaaga gagaaaaaac tttctggaac tgttgaggac tgcagatctg
6241 ccagtttggc tggcttacaa ggttgcagcg gctggagtgt cataccacga ccggaggtgg
6301 tgctttgatg gtcctaggac aaacacaatt ttagaagaca caacgaagt ggaagtcatc
6361 acgaagcttg gtgaaaggaa gattctgagg ccgcgctgga ttgacgccag ggtgtactcg
6421 gatcaccagg cactaaaggc gttcaaggac ttcgcctcgg gaaaacgttc tcagataggg
6481 ctcattgagg ttctgggaaa gatgcctgag cacttcatgg ggaagacatg ggaagcactt
6541 gacaccatgt acgttgtggc cactgcagag aaaggaggaa gagctcacag aatggccctg
6601 gaggaactgc cagatgctct tcagacaatt gccttgattg ccttattgac tgtgatgacc
6661 atgggggagtat tcttcctcct catgcagcgg aagggcattg gaaagatagg tttgggaggc
6721 gctgtcttgg gagtcgcgac cttttttctgt tggatggctg aagttccagg aacgaagatc
6781 gccggaatgt tgctgctctc ccttctcttg atgattgtgc taattcctga gccagagaag
6841 caacgttcgc agacagacaa ccagctagcc gtgttcctga tttgtgtcat gacccttgtg
6901 agcgcagtgg cagccaacga gatgggttgg ctagataaga ccaagagtga cataagcagt
6961 ttgtttgggc aaagaattga ggtcaaggag aatttcagca tgggagagtt tcttctggac
7021 ttgaggccgg caacagcctg gtcactgtac gctgtgacaa cagcggtcct cactccactg
7081 ctaaagcatt tgatcacgtc agattacatc aacacctcat tgacctcaat aaacgttcag
7141 gcaagtgcac tattcacact cgcgcgaggc ttcccctтcg tcgatgttgg agtgtcggct
7201 ctcctgctag cagccggatg ctggggacaa gtcaccctca ccgttacggt aacagcggca
7261 acactccttt tttgccacta tgcctacatg gttcccggtt ggcaagctga ggcaatgcgc
7321 tcagcccagc ggcggacagc ggccggaatc atgaagaacg ctgtagtgga tggcatcgtg
7381 gccacggacg tcccagaatt agagcgcacc acacccatca tgcagaagaa agttggacag
7441 atcatgctga tcttggtgtc tctagctgca gtagtagtga accgtctgt gaagacagta
7501 cgagaagccg gaattttgat cacggccgca gcggtgacgc tttgggagaa tggagcaagc
7561 tctgtttgga acgcaacaac tgccatcgga ctctgccaca tcatgcgtgg gggttggttg
7621 tcatgtctat ccataacatg gacactcata aagaacatgg aaaaaccagg actaaaaaga
7681 ggtggggcaa aaggacgcac cttgggagag gttttggaaag aaagactcaa ccagatgaca
7741 aaagaagagt tcactaggta ccgcaaagag gccatcatcg aagtcgatcg ctcagcggca
7801 aaacacgcca ggaaagaagg caatgtcact ggagggcatc cagtctctag gggcacagca
7861 aaactgagat ggctggtcga acggaggttt ctcgaaccgg tcggaaaagt gattgacctt
7921 ggatgtggaa gaggcggttg gtgttactat atggcaaccc aaaaaagagt ccaagaagtc
7981 agagggtaca caaagggcgg tcccggacat gaagagcccc aactagtgca aagttatgga
8041 tggaacattg tcaccatgaa gagtggagtg gatgtgttct acagaccttc tgagtgttgt
8101 gacactctcc tttgtgacat cggagagtcc tcgtcaagtg ctgaggttga agagcatagg
8161 acgattcggg tccttgaaat ggttgaggac tggctgcacc gagggccaag ggaattttgc
8221 gtgaaggtgc tctgccccta catgccgaaa gtcatagaga agatgaaggct gctccaacgc
8281 cggtatgggg ggggactggt cagaaaccca ctctcacgga attccacgca cgagatgtat
8341 tgggtgagtc gagcttcagg caatgtggta cattcagtga atatgaccag ccaggtgctc
8401 ctaggaagaa tggaaaaaag gacctggaag ggaccccaat acgaggaaga tgtaaacttg
8461 ggaagtggaa ccagggcggt gggaaaaccc ctgctcaact cagacaccag taaaatcaag
8521 aacaggattg aacgactcag gcgtgagtac agttcgacgt ggcaccacga tgagaaccac
8581 ccatatagaa cctggaacta tcacggcagt tatgatgtga gcccacagg ctccgccagt
8641 tcgctggtca atggagtggt caggctcctc tcaaaaccat gggacaccat cacgaatgtt
8701 accaccatgg ccatgactga cactactccc ttcgggcagc agcgagtgtt caaagagaag
8761 gtggacacga aagctcctga accgccagaa ggagtgaagt acgtgctcaa cgagaccacc
```

```
 8821 aactggttgt gggcgttttt ggccagagaa aaacgtccca gaatgtgctc tcgagaggaa
 8881 ttcataagaa aggtcaacag caatgcagct ttgggtgcca tgtttgaaga gcagaatcaa
 8941 tggaggagcg ccagagaagc agttgaagat ccaaaatttt gggagatggt ggatgaggag
 9001 cgcgaggcac atctgcgggg ggaatgtcac acttgcattt acaacatgat gggaaagaga
 9061 gagaaaaaac ccggagagtt cggaaaggcc aagggaagca gagccatttg gttcatgtgg
 9121 ctcggagctc gctttctgga gttcgaggct ctgggttttc tcaatgaaga ccactggctt
 9181 ggaagaaaga actcaggagg aggtgtcgag ggcttgggcc tccaaaaact gggttacatc
 9241 ctgcgtgaag ttggcacccg gcctgggggc aagatctatg ctgatgacac agctggctgg
 9301 gacacccgca tcacgagagc tgacttggaa aatgaagcta aggtgcttga gctgcttgat
 9361 ggggaacatc ggcgtcttgc cagggccatc attgagctca cctatcgtca caaagttgtg
 9421 aaagtgatgc gcccggctgc tgatggaaga accgtcatgg atgttatctc cagagaagat
 9481 cagaggggga gtggacaagt tgtcacctac gccctaaaca ctttcaccaa cctggccgtc
 9541 cagctggtga ggatgatgga aggggaagga gtgattggcc cagatgatgt ggagaaactc
 9601 acaaaaggga aaggacccaa agtcaggacc tggctgtttg agaatgggga agaaagactc
 9661 agccgcatgg ctgtcagtgg agatgactgt gtggtaaagc ccctggacga tcgctttgcc
 9721 acctcgctcc acttcctcaa tgctatgtca aaggttcgca aagacatcca agagtggaaa
 9781 ccgtcaactg gatggtatga ttggcagcag gttccatttt gctcaaacca tttcactgaa
 9841 ttgatcatga aagatggaag aacactggtg gttccatgcc gaggacagga tgaattggta
 9901 ggcagagctc gcatatctcc aggggccgga tggaacgtcc gcgacactgc ttgtctggct
 9961 aagtcttatg cccagatgtg gctgcttctg tacttccaca gaagagacct gcggctcatg
10021 gccaacgcca tttgctccgc tgtccctgtg aattgggtcc ctaccggaag aaccacgtgg
10081 tccatccatg caggaggaga gtggatgaca acagaggaca tgttggaggt ctggaaccgt
10141 gtttggatag aggagaatga atggatggaa gacaaaaccc cagtggagaa atggagtgac
10201 gtcccatatt caggaaaacg agaggacatc tggtgtggca gcctgattgg cacaagagcc
10261 cgagccacgt gggcagaaaa catccaggtg gctatcaacc aagtcagagc aatcatcgga
10321 gatgagaagt atgtggatta catgagttca ctaaagagat atgaagacac aactttggtt
10381 gaggacacga tactgtagat atttaatcaa ttgtaaatag acaatataag tatgcataaa
10441 agtgtagttt tatagtagta tttagtggtg ttagtgtaaa tagttaagaa aattttgagg
10501 agaaagtcag gccgggaagt tcccgccacc ggaagttgag tagacggtgc tgcctgcgac
10561 tcaaccccag gaggactggg tgaacaaagc cgcgaagtga tccatgtaag ccctcagaac
10621 cgtctcggaa ggaggacccc acatgttgta acttcaaagc ccaatgtcag accacgctac
10681 ggcgtgctac tctgcggaga gtgcagtctg cgatagtgcc ccaggaggac tgggttaaca
10741 aaggcaaacc aacgccccac gcggccctag ccccggtaat ggtgttaacc agggcgaaag
10801 gactagaggt tagaggagac cccgcggttt aaagtgcacg gcccagcctg gctgaagctg
10861 taggtcaggg gaaggactag aggttagtgg agaccccgtg ccacaaaaca ccacaacaaa
10921 acagcatatt gacacctggg atagactagg agatcttctg ctctgcacaa ccagccacac
10981 ggcacagtgc gccgacaatg gtggctggtg gtgcgagaac acaggatct
```
//

Apr 17 2007 11:10:07

My NCBI

[Sign In] [Register]

Nucleotide

Search CoreNucleotide for _____ [Go] [Clear]

Limits   Preview/Index   History   Clipboard   Details

Display GenBank   Show 5   Send to   Hide: ☐ sequence ☐ all but gene, CDS and mRNA features

Range: from begin ____ to end ____   ☐ Reverse complemented strand   Features: [+] [Refresh]

☐ 1: AF404755. Reports West Nile virus i...[gi:21929236]   Links

Features   Sequence

```
LOCUS       AF404755                11029 bp ss-RNA    linear   VRL 23-JUL-2002
DEFINITION  West Nile virus isolate WN NY 2000-grouse3282, complete genome.
ACCESSION   AF404755
VERSION     AF404755.1  GI:21929236
KEYWORDS    .
SOURCE      West Nile virus
  ORGANISM  West Nile virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
            Flavivirus; Japanese encephalitis virus group.
REFERENCE   1  (bases 1 to 11029)
  AUTHORS   Lanciotti,R.S., Ebel,G.D., Deubel,V., Kerst,A.J., Murri,S.,
            Meyer,R., Bowen,M., McKinney,N., Morrill,W.E., Crabtree,M.B.,
            Kramer,L.D. and Roehrig,J.T.
  TITLE     Complete genome sequences and phylogenetic analysis of West Nile
            virus strains isolated from the United States, Europe, and the
            Middle East
  JOURNAL   Virology 298 (1), 96-105 (2002)
   PUBMED   12093177
REFERENCE   2  (bases 1 to 11029)
  AUTHORS   Ebel,G.D., Kerst,A.J. and Lanciotti,R.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Division of Vector-Borne Infectious
            Diseases, Centers for Disease Control & Prevention, Rampart Road,
            Fort Collins, CO 80521, USA
FEATURES             Location/Qualifiers
     source          1..11029
                     /organism="West Nile virus"
                     /mol_type="genomic RNA"
                     /isolate="WN NY 2000-grouse3282"
                     /specific_host="ruffed grouse"
                     /db_xref="taxon:11082"
                     /country="USA: New York"
     CDS             97..10398
                     /note="contains capsid, pre-membrane, envelope, NS1, NS2a,
                     NS2b, NS3, NS4a, NS4b, and NS5"
                     /codon_start=1
                     /product="polyprotein precursor"
                     /protein_id="AAM81751.1"
                     /db_xref="GI:21929237"
                     /translation="MSKKPGGPGKSRAVNMLKRGMPRVLSLIGLKRAMLSLIDGKGPI
                     RFVLALLAFFRFTAIAPTRAVLDRWRGVNKQTAMKHLLSFKKELGTLTSAINRRSSKQ
                     KKRGGKTGIAVMIGLIASVGAVTLSNFQGKVMMTVNATDVTDVITIPTAAGKNLCIVR
                     AMDVGYMCDDTITYECPVLSAGNDPEDIDCWCTKSAVYVRYGRCTKTRHSRRSRRSLT
                     VQTHGESTLANKKGAWMDSTKATRYLVKTESWILRNPGYALVAAVIGWMLGSNTMQRV
                     VFVVLLLLVAPAYSFNCLGMSNRDFLEGVSGATWVDLVLEGDSCVTIMSKDKPTIDVK
                     MMNMEAANLAEVRSYCYLATVSDLSTKAACPTMGEAHNDKRADPAFVCRQGVVDRGWG
                     NGCGLFGKGSIDTCAKFACSTKAIGRTILKENIKYEVAIFVHGPTTVESHGNYSTQVG
                     ATQAGRFSITPAAPSYTLKLGEYGEVTVDCEPRSGIDTNAYYVMTVGTKTFLVHREWF
                     MDLNLPWSSAGSTVWRNRETLMEFEEPHATKQSVIALGSQEGALHQALAGAIPVEFSS
```

```
                        NTVKLTSGHLKCRVKMEKLQLKGTTYGVCSKAFKFLGTPADTGHGTVVLELQYTGTDG
                        PCKVPISSVASLNDLTPVGRLVTVNPFVSVATANAKVLIELEPPFGDSYIVVGRGEQQ
                        INHHWHKSGSSIGKAFTTTLKGAQRLAALGDTAWDFGSVGGVFTSVGKAVHQVFGGAF
                        RSLFGGMSWITQGLLGALLLWMGINARDRSIALTFLAVGGVLLFLSVNVHADTGCAID
                        ISRQELRCGSGVFIHNDVEAWMDRYKYPETPQGLAKIIQKAHKEGVCGLRSVSRLEH
                        QMWEAVKDELNTLLKENGVDLSVVVEKQEGMYKSAPKRLTATTEKLEIGWKAWGKSIL
                        FAPELANNTFVVDGPETKECPTQNRAWNSLEVEDFGFGLTSTRMFLKVRESNTTECDS
                        KIIGTAVKNNLAIHSDLSYWIESRLNDTWKLERAVLGEVKSCTWPETHTLWGDGILES
                        DLIIPVTLAGPRSNHNRRPGYKTQNQGPWDEGRVEIDFDYCPGTTVTLSESCGHRGPA
                        TRTTTESGKLITDWCCRSCTLPPLRYQTDSGCWYGMEIRPQRHDEKTLVQSQVNAYNA
                        DMIDPFQLGLLVVFLATQEVLRKRWTAKISMPAILIALLVLVFGGITYTDVLRYVILV
                        GAAFAESNSGGDVVHLALMATFKIQPVFMVASFLKARWTNQENILLMLAAVFFQMAYH
                        DARQILLWEIPDVLNSLAVAWMILRAITFTTTSNVVVPLLALLTPGLRCLNLDVYRIL
                        LLMVGIGSLIREKRSAAAKKKGASLLCLALASTGLFNPMILAAGLIACDPNRKRGWPA
                        TEVMTAVGLMFAIVGGLAELDIDSMAIPMTIAGLMFAAFVISGKSTDMWIERTADISW
                        ESDAEITGSSERVDVRLDDDGNFQLMNDPGAPWKIWMLRMVCLAISAYTPWAILPSVV
                        GFWITLQYTKRGGVLWDTPSPKEYKKGDTTTGVYRIMTRGLLGSYQAGAGVMVEGVFH
                        TLWHTTKGAALMSGEGRLDPYWGSVKEDRLCYGGPWKLQHKWNGQDEVQMIVVEPGKN
                        VKNVQTKPGVFKTPEGEIGAVTLDFPTGTSGSPIVDKNGDVIGLYGNGVIMPNGSYIS
                        AIVQGERMDEPIPAGFEPEMLRKKQITVLDLHPGAGKTRRILPQIIKEAINRRLRTAV
                        LAPTRVVAAEMAEALRGLPIRYQTSAVPREHNGNEIVDVMCHATLTHRLMSPHRVPNY
                        NLFVMDEAHFTDPASIAARGYISTKVELGEAAAIFMTATPPGTSDPFPESNSPISDLQ
                        TEIPDRAWNSGYEWITEYTGKTVWFVPSVKMGNEIALCLQRAGKKVVQLNRKSYETEY
                        PKCKNDDWDFVITTDISEMGANFKASRVIDSRKSVKPTIITEGEGRVILGEPSAVTAA
                        SAAQRRGRIGRNPSQVGDEYCYGGHTNEDDSNFAHWTEARIMLDNINMPNGLIAQFYQ
                        PEREKVYTMDGEYRLRGEERKNFLELLRTADLPVWLAYKVAAAGVSYHDRRWCFDGPR
                        TNTILEDNNEVEVITKLGERKILRPRWIDARVYSDHQALKAFKDFASGKRSQIGLIEV
                        LGKMPEHFMGKTWEALDTMYVVATAEKGGRAHRMALEELPDALQTIALIALLSVMTMG
                        VFFLLMQRKGIGKIGLGGAVLGVATFFCWMAEVPGTKIAGMLLLSLLLMIVLIPEPEK
                        QRSQTDNQLAVFLICVMTLVSAVAANEMGWLDKTKSDISSLFGQRIEVKENFSMGEFL
                        LDLRPATAWSLYAVTTAVLTPLLKHLITSDYINTSLTSINVQASALFTLARGFPFVDV
                        GVSALLLAAGCWGQVTLTVTVTAATLLFCHYAYMVPGWQAEAMRSAQRRTAAGIMKNA
                        VVDGIVATDVPELERTTPIMQKKVGQIMLILVSLAAVVVNPSVKTVREAGILITAAAV
                        TLWENGASSVWNATTAIGLCHIMRGGWLSCLSITWTLIKNMEKPGLKRGGAKGRTLGE
                        VWKERLNQMTKEEFTRYRKEAIIEVDRSAAKHARKEGNVTGGHPVSRGTAKLRWLVER
                        RFLEPVGKVIDLGCGRGGWCYYMATQKRVQEVRGYTKGGPGHEEPQLVQSYGWNIVTM
                        KSGVDVFYRPSECCDTLLCDIGESSSSAEVEEHRTIRVLEMVEDWLHRGPREFCVKVL
                        CPYMPKVIEKMELLQRRYGGGLVRNPLSRNSTHEMYWVSRASGNVVHSVNMTSQVLLG
                        RMEKRTWKGPQYEEDVNLGSGTRAVGKPLLNSDTSKIKNRIERLRREYSSTWHHDENH
                        PYRTWNYHGSYDVKPTGSASSLVNGVVRLLSKPWDTITNVTTMAMTDTTPFGQQRVFK
                        EKVDTKAPEPPEGVKYVLNETTNWLWAFLAREKRPRMCSREEFIRKVNSNAALGAMFE
                        EQNQWRSAREAVEDPKFWEMVDEEREAHLRGECHTCIYNMMGKREKKPGEFGKAKGSR
                        AIWFMWLGARFLEFEALGFLNEDHWLGRKNSGGGVEGLGLQKLGYILREVGTRPGGKI
                        YADDTAGWDTRITRADLENEAKVLELLDGEHRRLARAIIELTYRHKVVKVMRPAADGR
                        TVMDVISREDQRGSGQVVTYALNTFTNLAVQLVRMMEGEGVIGPDDVEKLTKGKGPKV
                        RTWLFENGEERLSRMAVSGDDCVVKPLDDRFATSLHFLNAMSKVRKDIQEWKPSTGWY
                        DWQQVPFCSNHFTELIMKDGRTLVVPCRGQDELVGRARISPGAGWNVRDTACLAKSYA
                        QMWLLLYFHRRDLRLMANAICSAVPVNWVPTGRTTWSIHAGGEWMTTEDMLEVWNRVW
                        IEENEWMEDKTPVEKWSDVPYSGKREDIWCGSLIGTRARATWAENIQVAINQVRAIIG
                        DEKYVDYMSSLKRYEDTTLVEDTVL"
ORIGIN
        1 agtagttcgc ctgtgtgagc tgacaaactt agtagtgttt gtgaggatta acaacaatta
       61 acacagtgcg agctgtttct tagcacgaag atttcgatgt ctaagaaacc aggagggccc
      121 ggcaagagcc gggctgtcaa tatgctaaaa cgcggaatgc cccgcgtgtt gtccttgatt
      181 ggactgaaga gggctatgtt gagcctgatc gacggcaagg ggccaatacg atttgtgttg
      241 gctctcttgg cgttcttcag gttcacagca attgctccga cccgagcagt gctggatcga
      301 tggagaggtg tgaacaaaca aacagcgatg aaacaccttc tgagttttaa gaaggaacta
      361 gggaccttga ccagtgctat caatcggcgg agctcaaaac aaaagaaaag aggaggaaag
      421 accggaattg cagtcatgat tggcctgatc gccagcgtag gagcagttac cctctctaac
      481 ttccaaggga aggtgatgat gacggtaaat gctactgacg tcacagatgt catcacgatt
      541 ccaacagctg ctggaaagaa cctatgcatt gtcagagcaa tggatgtggg atacatgtgc
      601 gatgatacta tcacttatga atgcccagtg ctgtcggctg gtaatgatcc agaagacatc
      661 gactgttggt gcacaaagtc agcagtctac gtcaggtatg gaagatgcac aagacacgc
      721 cactcaagac gcagtcggag gtcactgaca gtgcagacac acggagaaag cactctagcg
      781 aacaagaagg gggcttggat ggacagcacc aaggccacaa ggtatttggt aaaaacagaa
      841 tcatggatct tgaggaaccc tggatatgcc ctggtggcag ccgtcattgg ttggatgctt
```

```
 901 gggagcaaca ccatgcagag agttgtgttt gtcgtgctat tgcttttggt ggccccagct
 961 tacagcttca actgccttgg aatgagcaac agagacttct tggaaggagt gtctggagca
1021 acatgggtgg atttggttct cgaaggcgac agctgcgtga ctatcatgtc taaggacaag
1081 cctaccatcg atgtgaagat gatgaatatg gaggcggcca acctggcaga ggtccgcagt
1141 tattgctatt tggctaccgt cagcgatctc tccaccaaag ctgcgtgccc gaccatggga
1201 gaagctcaca atgacaaacg tgctgaccca gcttttgtgt gcagacaagg agtggtggac
1261 aggggctggg gcaacggctg cggattattt ggcaaaggaa gcattgacac atgcgccaaa
1321 tttgcctgct ctaccaaggc aataggaaga accatcttga aagagaatat caagtacgaa
1381 gtggccattt ttgtccatgg accaactact gtggagtcgc acggaaacta ctccacacag
1441 gttggagcca ctcaggcagg gagattcagc atcactcctg cggcgccttc atacacacta
1501 aagcttggag aatatggaga ggtgacagtg gactgtgaac cacggtcagg gattgacacc
1561 aatgcatact acgtgatgac tgttggaaca aagacgttct tggtccatcg tgagtggttc
1621 atggacctca acctcccttg gagcagtgct ggaagtactg tgtggaggaa cagagagacg
1681 ttaatggagt ttgaggaacc acacgccacg aagcagtctg tgatagcatt gggctcacaa
1741 gagggagctc tgcatcaagc tttggctgga gccattcctg tggaattttc aagcaacact
1801 gtcaagttga cgtcgggtca tttgaagtgt agagtgaaga tggaaaaatt gcagttgaag
1861 ggaacaacct atggcgtctg ttcaaaggct ttcaagtttc ttgggactcc cgcagacaca
1921 ggtcacggca ctgtggtgtt ggaattgcag tacactggca cggatggacc ttgcaaagtt
1981 cctatctcgt cagtggcttc attgaacgac ctaacgccag tgggcagatt ggtcactgtc
2041 aacccttttg tttcagtggc cacggccaac gctaaggtcc tgattgaatt ggaaccaccc
2101 tttggagact catacatagt ggtgggcaga ggagaacaac agatcaatca ccattggcac
2161 aagtctggaa gcagcattgg caaagccttt acaaccaccc tcaaaggagc gcagagacta
2221 gccgctctag gagacacagc ttgggacttt ggatcagttg gaggggtgtt cacctcagtt
2281 gggaaggctg tccatcaagt gttcggagga gcattccgct cactgttcgg aggcatgtcc
2341 tggataacgc aaggattgct gggggctctc ctgttgtgga tgggcatcaa tgctcgtgat
2401 aggtccatag ctctcacgtt tctcgcagtt ggaggagttc tgctcttcct ctccgtgaac
2461 gtgcacgctc acactgggtg tgccatagac atcagccggc aagagctgag atgtggaagt
2521 ggagtgttca tacacaatga tgtggaggct tggatggacc ggtacaagta ttaccctgaa
2581 acgccacaag gcctagccaa gatcattcag aaagctcata aggaaggagt gtgcggtcta
2641 cgatcagttt ccagactgga gcatcaaatg tgggaagcag tgaaggacga gctgaacact
2701 cttttgaagg agaatggtgt ggaccttagt gtcgtggttg agaaacagga gggaatgtac
2761 aagtcagcac ctaaacgcct caccgccacc acggaaaaat tggaaattgg ctggaaggcc
2821 tggggaaaga gtattttatt tgcaccagaa ctcgccaaca cacctttgt ggttgatggt
2881 ccggagacca aggaatgtcc gactcagaat cgcgcttgga atagcttaga agtggaggat
2941 tttggatttg gtctcaccag cactcggatg ttcctgaagg tcagagagag caacacaact
3001 gaatgtgact cgaagatcat tggaacggct gtcaagaaca acttggcgat ccacagtgac
3061 ctgtcctatt ggattgaaag caggctcaat gatacgtgga agcttgaaag ggcagttctg
3121 ggtgaagtca aatcatgtac gtggcctgag acgcatacct tgtggggcga tggaatcctt
3181 gagagtgact tgataatacc agtcacactg gcgggaccac gaagcaatca caatcggaga
3241 cctgggtaca agacacaaaa ccagggccca tgggacgaag gccgggtaga gattgacttc
3301 gattactgcc caggaactac ggtcaccctg agtgagagct gcggacaccg tggacctgcc
3361 actcgcacca ccacagagag cggaaagttg ataacagatt ggtgctgacg gagctgcacc
3421 ttaccaccac tgcgctacca aactgacagc ggctgttggt atggtatgga gatcagacca
3481 cagagacatg atgaaaagac cctcgtgcag tcacaagtga atgcttataa tgctgatatg
3541 attgaccctt ttcagttggg ccttctggtc gtgttcttgg ccacccagga ggtccttcgc
3601 aagaggtgga cagccaagat cagcatgcca gctatactga ttgctctgct agtcctggtg
3661 tttgggggca ttacttacac tgatgtgtta cgctatgtca tcttggtggg ggcagctttc
3721 gcagaatcta attcgggagg agacgtggta cacttggcgc tcatggcgac cttcaagata
3781 caaccagtgt ttatggtggc atcgtttctc aaagcgagat ggaccaacca ggagaacatt
3841 ttgttgatgt tggcggctgt tttctttcaa atggcttatc acgatgcccg ccaaattctg
3901 ctctgggaga tccctgatgt gttgaattca ctggcggtag cttggatgat actgagagcc
3961 ataacattca caacgacatc aaacgtggtt gttccgctgc tagccctgct aacacccggg
4021 ctgagatgct tgaatctgga tgtgtacagg atactgctgt gatggtcgg aataggcagc
4081 ttgatcaggg agaagaggag tgcagctgca aaaagaaag gagcaagtct gctatgcttg
4141 gctctggcct caacaggact tttcaacccc atgatccttg ctgctggact gattgcatgt
4201 gatcccaacc gtaaacgcgg atggcccgca actgaagtga tgacagctgt cggcctaatg
4261 tttgccatcg tcggagggct ggcagagctt gacattgact ccatgccat tccaatgact
4321 atcgcggggc tcatgtttgc tgctttcgtg atttctggga aatcaacaga tatgtggatt
4381 gagagaacgg cggacatttc ctgggaaagt gatgcagaaa ttacaggctc gagcgaaaga
4441 gttgatgtgc ggcttgatga tgatggaaac ttccagctca tgaatgatcc aggagcacct
4501 tggaagatat ggatgctcag aatggtctgt ctcgcgatta gtgcgtacac ccctgggca
4561 atcctgccct cagtagttgg attttggata actctccaat acacaaagag aggaggcgtg
4621 ttgtgggaca ctcccctcacc aaaggagtac aaaaaggggg acacgaccac cggcgtctac
4681 aggatcatga ctcgtgggct gctcggcagt tatcaagcag gagcgggcgt gatggttgaa
4741 ggtgtttttc acaccctttg gcatacaaca aaaggagccg ctttgatgag cggagagggc
4801 cgcctggacc catactgggg cagtgtcaag gaggatcgac tttgttacgg aggaccctgg
```

```
4861 aaattgcagc acaagtggaa cgggcaggat gaggtgcaga tgattgtggt ggaacctggc
4921 aagaacgtta agaacgtcca gacgaaacca ggggtgttca aaacacctga aggagaaatc
4981 ggggccgtga ctttggactt ccccactgga acatcaggct caccaatagt ggacaaaaac
5041 ggtgatgtga ttgggcttta tggcaatgga gtcataatgc ccaacggctc atacataagc
5101 gcgatagtgc agggtgaaag gatggatgag ccaatcccag ccggattcga acctgagatg
5161 ctgaggaaaa aacagatcac tgtactggat ctccatcccg gcgccggtaa aacaaggagg
5221 attctgccac agatcatcaa agaggccata aacagaagac tgagaacagc cgtgctagca
5281 ccaaccaggg ttgtggctgc tgagatggct gaagcactga gaggactgcc catccggtac
5341 cagacatccg cagtgcccag agaacataat ggaaatgaga ttgttgatgt catgtgtcat
5401 gctaccctca cccacaggct gatgtctcct cacagggtgc cgaactacaa cctgttcgtg
5461 atggatgagg ctcatttcac cgacccagct agcattgcag caagaggtta catttccaca
5521 aaggtcgagc taggggaggc ggcggcaata ttcatgacag ccacccccac aggcacttca
5581 gatccattcc cagagtccaa ttcaccaatt tccgacttac agactgagat cccggatcga
5641 gcttggaact ctggatacga atggtcaca gaatacaacg ggaagacggt ttggtttgtg
5701 cctagtgtca agatgggga tgagattgcc ctttgcctac aacgtgctgg aaagaaagta
5761 gtccaattga acagaaagtc gtacgagacg gagtacccaa aatgtaagaa cgatgattgg
5821 gactttgtta tcacaacaga catatctgaa atgggggcta acttcaaggc gagcagggtg
5881 attgacagcc ggaagagtgt gaaaccaacc atcataacag aaggagaagg gagagtgatc
5941 ctgggagaac catctgcagt gacagcagct agtgccgccc agagacgtgg acgtatcggt
6001 agaaatccgt cgcaagttgg tgatgagtac tgttatgggg ggcacacgaa tgaagacgac
6061 tcgaacttcg cccattggac tgaggcacga atcatgctgg acaacatcaa catgccaaac
6121 ggactgatcg ctcaatttta ccaaccagag cgtgagaagg tatataccat ggatggggaa
6181 taccggctca gaggagaaga gagaaaaaac tttctggaac tgttgaggac tgcagatctg
6241 ccagtttggc tggcttacaa ggttgcagcg gctggagtgt cataccacga ccggaggtgg
6301 tgctttgatg gtcctaggac aaacacaatt ttagaagaca caacgaagt ggaagtcatc
6361 acgaagcttg gtgaaaggaa gattctgagg ccgcgctgga ttgacgccag ggtgtactcg
6421 gatcatcagg cactaaaggc gttcaaggac ttcgcctcgg gaaaacgttc tcagataggg
6481 ctcattgagg ttctgggaaa gatgcctgag cacttcatgg ggaagacatg ggaagcactt
6541 gacaccatgt acgttgtggc cactgcagga aaaggaggaa gagctcacag aatggccctg
6601 gaggaactgc cagatgctct tcagacaatt gccttgattg ccttattgag tgtgatgacc
6661 atgggagtat tcttcctcct catgcagcgg aagggcattg gaaagatagg tttgggaggc
6721 gctgtcttgg gagtcgcgac cttttttctgt tggatggctg aagttccagg acgaagatc
6781 gccggaatgt tgctgctctc ccttctcttg atgattgtgc taattcctga gccagagaag
6841 caacgttcgc agacagacaa ccagctagcc gtgttcctga tttgtgtcat gacccttgtg
6901 agcgcagtgg cagccaacga gatgggttgg ctagataaga ccaagagtga cataagcagt
6961 ttgtttgggc aaagaattga ggtcaaggag aatttcagca tgggagagtt tcttctggac
7021 ttgaggccgg caacagcctg gtcactgtac gctgtgacaa cagcggtcct cactccactg
7081 ctaaagcatt tgatcacgtc agattacatc aacacctcat tgacctcaat aaacgttcag
7141 gcaagtgcac tattcacact cgcgcgaggc ttccccttcg tcgatgttgg agtgtcggct
7201 ctcctgctag cagccggatg ctgggggacaa gtcaccctca ccgttacggt aacagcggca
7261 acactccttt tttgccacta tgcctacatg gttcccggtt ggcaagctga ggcaatgcgc
7321 tcagcccagc ggcggaacagc ggccggaatc atgaagaacg ctgtagtgga tggcatcgtg
7381 gccacggacg tcccagaatt agagcgcacc acacccatca tgcagaagaa agttggacag
7441 atcatgctga tcttggtgtc tctagctgca gtagtagtga acccgtctgt gaagacagta
7501 cgagaagccg gaattttgat cacggccgca gcggtgacgc tttgggagaa tggagcaagc
7561 tctgtttgga acgcaacaac tgccatcgga ctctgccaca tcatgcgtgg gggttggttg
7621 tcatgtctat ccataacatg gacactcata aagaacatgg aaaaaccagg actaaaaaga
7681 ggtggggcaa aaggacgcac cttgggagag gttttggaaag aaagactcaa ccagatgaca
7741 aaagaagagt tcactaggta ccgcaaagag gccatcatcg aagtcgatcg ctcagcggca
7801 aaacacgcca ggaaagaagg caatgtcact ggagggcatc cagtctctag gggcacagca
7861 aaactgagat ggctggtcga acggaggttt ctcgaaccgg tcggaaaagt gattgacctt
7921 ggatgtggaa gaggcggctg gtgttactat atggcaaccc aaaaaagagt ccaagaagtc
7981 agagggtaca caaagggcgg tcccggacat gaagagcccc aactagtgca aagttatgga
8041 tggaacattg tcaccatgaa gagtggggtg gatgtgttct acagaccttc tgagtgttgt
8101 gacaccctcc tttgtgacat cggagagtcc tcgtcaagtg ctgaggttga agagcatagg
8161 acgattcggg tccttgaaat ggttgaggac tggctcgacc gagggccaag ggaattttgc
8221 gtgaaggtgc tctgccccta catgccgaaa gtcatagaga agatggagtt gctccaacgc
8281 cggtatgggg ggggactggt cagaaaccca ctctcacgga attccacgca cgagatgtat
8341 tgggtgagtc gagcttcagg caatgtggta cattcagtga atatgaccag ccaggtgctc
8401 ctaggaagaa tggaaaaaag gacctggaag ggaccccaat acgaggaaga tgtaaacttg
8461 ggaagtggaa ccagggcggt gggaaaaccc ctgctcaact cagacaccag taaaatcaag
8521 aacaggattg aacgactcag gcgtgagtac agttcgacgt ggcaccacga tgagaaccac
8581 ccatatagaa cctggaacta tcacggcagt tatgatgtga gcccacagg ctccgccagt
8641 tcgctggtca tggagtggt caggctcctc tcaaaaccat gggacaccat cacgaatgtt
8701 accaccatgg ccatgactga cactactccc ttcgggcagc agcgagtgtt caaagagaag
8761 gtggacacga agctcctga accgccagaa ggagtgaagt acgtgctcaa cgagaccacc
```
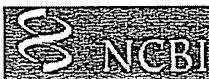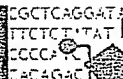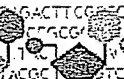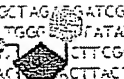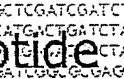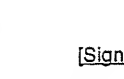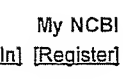
```
 8821 aactggttgt gggcgttttt ggccagagaa aaacgtccca gaatgtgctc tcgagaggaa
 8881 ttcataagaa aggtcaacag caatgcagct ttgggtgcca tgtttgaaga gcagaatcaa
 8941 tggaggagcg ccagagaagc agttgaagat ccaaaatttt gggagatggt ggatgaggag
 9001 cgcgaggcac atctgcgggg ggaatgtcac acttgcattt acaacatgat gggaaagaga
 9061 gagaaaaaac ccggagagtt cggaaaggcc aagggaagca gagccatttg gttcatgtgg
 9121 ctcggagctc gctttctgga gttcgaggct ctgggttttc tcaatgaaga ccactggctt
 9181 ggaagaaaga actcaggagg aggtgtcgag ggcttgggcc tccaaaaact gggttacatc
 9241 ctgcgtgaag ttggcacccg gcctgggggc aagatctatg ctgatgacac agctggctgg
 9301 gacacccgca tcacgagagc tgacttggaa aatgaagcta aggtgcttga gctgcttgat
 9361 ggggaacatc ggcgtcttgc cagggccatc attgagctca cctatcgtca caaagttgtg
 9421 aaagtgatgc gcccggctgc tgatggaaga accgtcatgg atgttatctc cagagaagat
 9481 cagaggggga gtggacaagt tgtcacctac gccctaaaca ctttcaccaa cctggccgtc
 9541 cagctggtga ggatgatgga aggggaagga gtgattggcc cagatgatgt ggagaaactc
 9601 acaaaaggga aaggacccaa agtcaggacc tggctgtttg agaatgggga agaaagactc
 9661 agccgcatgg ctgtcagtgg agatgactgt gtggtaaagc ccctggacga tcgctttgcc
 9721 acctcgctcc acttcctcaa tgctatgtca aaggttcgca aagacatcca agagtggaaa
 9781 ccgtcaactg gatggtatga ttggcagcag gttccatttt gctcaaacca tttcactgaa
 9841 ttgatcatga aagatggaag aacactggtg gttccatgcc gaggacagga tgaattggta
 9901 ggcagagctc gcatatctcc aggggccgga tggaacgtcc gcgacactgc ttgtctggct
 9961 aagtcttatg cccagatgtg gctgcttctg tacttccaca gaagagacct gcggctcatg
10021 gccaacgcca tttgctccgc tgtccctgtg aattgggtcc ctaccggaag aaccacgtgg
10081 tccatccatg caggaggaga gtggatgaca acagaggaca tgttggaggt ctggaaccgt
10141 gttttggatag aggagaatga atggatggaa gacaaaaccc cagtggagaa atggagtgac
10201 gtcccatatt caggaaaacg agaggacatc tggtgtggca gcctgattgg cacaagagcc
10261 cgagccacgt gggcagaaaa catccaggtg gctatcaacc aagtcagagc aatcatcgga
10321 gatgagaagt atgtggatta catgagttca ctaaagagat atgaagacac aactttggtt
10381 gaggacacag tattgtagat atttaatcaa ttgtaaatag acaatataag tatgcataaa
10441 agtgtagttt tatagtagta tttagtggtg ttagtgtaaa tagttaagaa aattttgagg
10501 agaaagtcag gccgggaagt tcccgccacc ggaagttgag tagacggtgc tgcctgcgac
10561 tcaaccccag gaggactggg tgaacaaagc cgcgaagtga tccatgtaag ccctcagaac
10621 cgtctcggaa ggaggacccc acatgttgta acttcaaagc ccaatgtcag accacgctac
10681 ggcgtgctac tctgcggaga gtgcagtctg cgatagtgcc ccaggaggac tgggttaaca
10741 aaggcaaacc aacgccccac gcggccctag ccccggtaat ggtgttaacc agggcgaaag
10801 gactagaggt tagaggagac cccgcggttt aaagtgcacg gcccagcctg gctgaagctg
10861 taggtcaggg gaaggactag aggttagtgg agaccccgtg ccacaaaaca ccacaacaaa
10921 acagcatatt gacacctggg atagactagg agatcttctg ctctgcacaa ccagccacac
10981 ggcacagtgc gccgacaatg gtggctggtg gtgcgagaac acaggatct
```
//

Apr 17 2007 11:10:07

Nucleotide

PubMed   Nucleotide   Protein   Genome   Structure   PMC   Taxonomy   OMIM   Books

Search : CoreNucleotide   for   [Go] [Clear]

Limits   Preview/Index   History   Clipboard   Details

Display GenBank   Show 5   Send to   Hide: ☐ sequence ☐ all but gene, CDS and mRNA features

Range: from begin   to end   ☐ Reverse complemented strand Features: [+] [Refresh]

☐ **1:** AF404756. Reports West Nile virus i...[gi:21929238]   Links

Features   Sequence

```
LOCUS       AF404756               11029 bp ss-RNA     linear    VRL 23-JUL-2002
DEFINITION  West Nile virus isolate WN NY 2000-crow3356, complete genome.
ACCESSION   AF404756
VERSION     AF404756.1  GI:21929238
KEYWORDS    .
SOURCE      West Nile virus
  ORGANISM  West Nile virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
            Flavivirus; Japanese encephalitis virus group.
REFERENCE   1  (bases 1 to 11029)
  AUTHORS   Lanciotti,R.S., Ebel,G.D., Deubel,V., Kerst,A.J., Murri,S.,
            Meyer,R., Bowen,M., McKinney,N., Morrill,W.E., Crabtree,M.B.,
            Kramer,L.D. and Roehrig,J.T.
  TITLE     Complete genome sequences and phylogenetic analysis of West Nile
            virus strains isolated from the United States, Europe, and the
            Middle East
  JOURNAL   Virology 298 (1), 96-105 (2002)
  PUBMED    12093177
REFERENCE   2  (bases 1 to 11029)
  AUTHORS   Ebel,G.D., Kerst,A.J. and Lanciotti,R.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Division of Vector-Borne Infectious
            Diseases, Centers for Disease Control & Prevention, Rampart Road,
            Fort Collins, CO 80521, USA
FEATURES             Location/Qualifiers
     source          1..11029
                     /organism="West Nile virus"
                     /mol_type="genomic RNA"
                     /isolate="WN NY 2000-crow3356"
                     /specific_host="crow"
                     /db_xref="taxon:11082"
                     /country="USA: New York"
     CDS             97..10398
                     /note="contains capsid, pre-membrane, envelope, NS1, NS2a,
                     NS2b, NS3, NS4a, NS4b, and NS5"
                     /codon_start=1
                     /product="polyprotein precursor"
                     /protein_id="AAM81752.1"
                     /db_xref="GI:21929239"
                     /translation="MSKKPGGPGKSRAVNMLKRGMPRVLSLIGLKRAMLSLIDGKGPI
                     RFVLALLAFFRFTAIAPTRAVLDRWRGVNKQTAMKHLLSFKKELGTLTSAINRRSSKQ
                     KKRGGKTGIAVMIGLIASVGAVTLSNFQGKVMMTVNATDVTDVITIPTAAGKNLCIVR
                     AMDVGYMCDDTITYECPVLSAGNDPEDIDCWCTKSAVYVRYGRCTKTRHSRRSRRSLT
                     VQTHGESTLANKKGAWMDSTKATRYLVKTESWILRNPGYALVAAVIGWMLGSNTMQRV
                     VFVVLLLLVAPAYSFNCLGMSNRDFLEGVSGATWVDLVLEGDSCVTIMSKDKPTIDVK
                     MMNMEAANLAEVRSYCYLATVSDLSTKAACPTMGEAHNDKRADPAFVCRQGVVDRGWG
                     NGCGLFGKGSIDTCAKFACSTKAIGRTILKENIKYEVAIFVHGPTTVESHGNYSTQVG
                     ATQAGRFSITPAAPSYTLKLGEYGEVTVDCEPRSGIDTNAYYVMTVGTKTFLVHREWF
                     MDLNLPWSSAGSTVWRNRETLMEFEEPHATKQSVIALGSQEGALHQALAGAIPVEFSS
```

```
                      NTVKLTSGHLKCRVKMEKLQLKGTTYGVCSKAFKFLGTPADTGHGTVVLELQYTGTDG
                      PCKVPISSVASLNDLTPVGRLVTVNPFVSVATANAKVLIELEPPFGDSYIVVGRGEQQ
                      INHHWHKSGSSIGKAFTTTLKGAQRLAALGDTAWDFGSVGGVFTSVGKAVHQVFGGAF
                      RLLFGGMSWITQGLLGALLLWMGINARDRSIALTFLAVGGVLLFLSVNVHADTGCAID
                      ISRQELRCGSGVFIHNDVEAWMDRYKYYPETPQGLAKIIQKAHKEGVCGLRSVSRLEH
                      QMWEAVKDELNTLLKENGVDLSVVVEKQEGMYKSAPKRLTATTEKLEIGWKAWGKSIL
                      FAPELANNTFVVDGPETKECPTQNRAWNSLEVEDFGFGLTSTRMFLKVRESNTTECDS
                      KIIGTAVKNNLAIHSDLSYWIESRLNDTWKLERAVLGEVKSCTWPETHTLWGDGILES
                      DLIIPVTLAGPRSNHNRRPGYKTQNQGPWDEGRVEIDFDYCPGTTVTLSESCGHRGPA
                      TRTTTESGKLITDWCCRSCTLPPLRYQTDSGCWYGMEIRPQRHDEKTLVQSQVNAYNA
                      DMIDPFQLGLLVVFLATQEVLRKRWTAKISMPAILIALLVLVFGGITYTDVLRYVILV
                      GAAFAESNSGGDVVHLALMATFKIQPVFMVASFLKARWTNQENILLMLAAVFFQMAYH
                      DARQILLWEIPDVLNSLAVAWMILRAITFTTTSNVVVPLLALLTPGLRCLNLDVYRIL
                      LLMVGIGSLIREKRSAAAKKKGASLLCLALASTGLFNPMILAAGLITCDPNRKRGWPA
                      TEVMTAVGLMFAIVGGLAELDIDSMAIPMTIAGLMFAAFVISGKSTDMWIERTADISW
                      ESDAEITGSSERVDVRLDDDGNFQLMNDPGAPWKIWMLRMVCLAISAYTPWAILPSVV
                      GFWITLQYTKRGGVLWDTPSPKEYKKGDTTTGVYRIMTRGLLGSYQAGAGVMVEGVFH
                      TLWHTTKGAALMSGEGRLDPYWGSVKEDRLCYGGPWKLQHKWNGQDEVQMIVVEPGKN
                      VKNVQTKPGVFKTPEGEIGAVTLDFPTGTSGSPIVDKNGDVIGLYGNGVIMPNGSYIS
                      AIVQGERMDEPIPAGFEPEMLRKKQITVLDLHPGAGKTRRILPQIIKEAINRRLRTAV
                      LAPTRVVAAEMAEALRGLPIRYQTSAVPREHNGNEIVDVMCHATLTHRLMSPHRVPNY
                      NLFVMDEAHFTDPASIAARGYISTKVELGEAAAIFMTATPPGTSDPFPESNSPISDLQ
                      TEIPDRAWNSGYEWITEYTGKTVWFVPSVKMGNEIALCLQRAGKKVVQLNRKSYETEY
                      PKCKNDDWDFVITTDISEMGANFKASRVIDSRKSVKPTIITEGEGRVILGEPSAVTAA
                      SAAQRRGRIGRNPSQVGDEYCYGGHTNEDDSNFAHWTEARIMLDNINMPNGLIAQFYQ
                      PEREKVYTMDGEYRLRGEERKNFLELLRTADLPVWLAYKVAAAGVSYHDRRWCFDGPR
                      TNTILEDNNEVEVITKLGERKILRPRWIDARVYSDHQALKAFKDFASGKRSQIGLIEV
                      LGKMPEHFMGKTWEALDTMYVVATAEKGGRAHRMALEELPDALQTIALIALLSVMTMG
                      VFFLLMQRKGIGKIGLGGAVLGVATFFCWMAEVPGTKIAGMLLLSLLLMIVLIPEPEK
                      QRSQTDNQLAVFLICVMTLVSAVAANEMGWLDKTKSDISSLFGQRIEVKENFSMGEFL
                      LDLRPATAWSLYAVTTAVLTPLLKHLITSDYINTSLTSINVQASALFTLARGFPFVDV
                      GVSALLLAAGCWGQVTLTVTVTAATLLFCHYAYMVPGWQAEAMRSAQRRTAAGIMKNA
                      VVDGIVATDVPELERTTPIMQKKVGQIMLILVSLAAVVVNPSVKTVREAGILITAAAV
                      TLWENGASSVWNATTAIGLCHIMRGGWLSCLSITWTLIKNMEKPGLKRGGAKGRTLGE
                      VWKERLNQMTKEEFTRYRKEAIIEVDRSAAKHARKEGNVTGGHPVSRGTAKLRWLVER
                      RFLEPVGKVIDLGCGRGGWCYYMATQKRVQEVRGYTKGGPGHEEPQLVQSYGWNIVTM
                      KSGVDVFYRPSECCDTLLCDIGESSSSAEVEEHRTIRVLEMVEDWLHRGPREFCVKVL
                      CPYMPKVIEKMELLQRRYGGGLVRNPLSRNSTHEMYWVSRASGNVVHSVNMTSQVLLG
                      RMEKRTWKGPQYEEDVNLGSGTRAVGKPLLNSDTSKIKNRIERLRREYSSTWHHDENH
                      PYRTWNYHGSYDVKPTGSASSLVNGVVRLLSKPWDTITNVTTMAMTDTTPFGQQRVFK
                      EKVDTKAPEPPEGVKYVLNETTNWLWAFLAREKRPRMCSREEFIRKVNSNAALGAMFE
                      EQNQWRSAREAVEDPKFWEMVDEEREAHLRGECHTCIYNMMGKREKKPGEFGKAKGSR
                      AIWFMWLGARFLEFEALGFLNEDHWLGRKNSGGGVEGLGLQKLGYILREVGTRPGGKI
                      YADDTAGWDTRITRADLENEAKVLELLDGEHRRLARAIIELTYRHKVVKVMRPAADGR
                      TVMDVISREDQRGSGQVVTYALNTFTNLAVQLVRMMEGEGVIGPDDVEKLTKGKGPKV
                      RTWLFENGEERLSRMAVSGDDCVVKPLDDRFATSLHFLNAMSKVRKDIQEWKPSTGWY
                      DWQQVPFCSNHFTELIMKDGRTLVVPCRGQDELVGRARISPGAGWNVRDTACLAKSYA
                      QMWLLLYFHRRDLRLMANAICSAVPVNWVPTGRTTWSIHAGGEWMTTEDMLEVWNRVW
                      IEENEWMEDKTPVEKWSDVPYSGKREDIWCGSLIGTRARATWAENIQVAINQVRAIIG
                      DEKYVDYMSSLKRYEDTTLVEDTVL"
          ORIGIN
                1 agtagttcgc ctgtgtgagc tgacaaactt agtagtgttt gtgaggatta acaacaatta
               61 acacagtgcg agctgtttct tagcacgaag atctcgatgt ctaagaaacc aggagggccc
              121 ggcaagagcc gggctgtcaa tatgctaaaa cgcggaatgc cccgcgtgtt gtccttgatt
              181 ggactgaaga gggctatgtt gagcctgatc gacggcaagg ggccaatacg atttgtgttg
              241 gctctcttgg cgttcttcag gttcacagca attgctccga cccgagcagt gctggatcga
              301 tggagaggtg tgaacaaaca aacagcgatg aaacaccttc tgagttttaa gaaggaacta
              361 gggaccttga ccagtgctat caatcggcgg agctcaaaac aaaagaaaag aggaggaaag
              421 accggaattg cagtcatgat tggcctgatc gccagcgtag gagcagttac cctctctaac
              481 ttccaaggga aggtgatgat gacggtaaat gctactgacg tcacagatgt catcacgatt
              541 ccaacagctg ctggaaagaa cctatgcatt gtcagagcaa tggatgtggg atacatgtgc
              601 gatgatacta tcacttatga atgcccagtg ctgtcggctg taatgatcc agaagacatc
              661 gactgttggt cacaaagtc agcagctac gtcaggtatg aagatgcac caagacacgc
              721 cactcaagac gcagtcggag gtcactgaca gtgcagacac acggagaaag cactctagcg
              781 aacaagaagg gggcttggat ggacagcacc aaggccacaa ggtatttggt aaaaacagaa
              841 tcatggatct tgaggaaccc tggatatgcc ctggtggcag ccgtcattgg ttggatgctt
```

```
 901 gggagcaaca ccatgcagag agttgtgttt gtcgtgctat tgctttggt ggccccagct
 961 tacagcttca actgccttgg aatgagcaac agagacttct tggaaggagt gtctggagca
1021 acatgggtgg atttggttct cgaaggcgac agctgcgtga ctatcatgtc taaggacaag
1081 cctaccatcg atgtgaagat gatgaatatg gaggcggcca acctggcaga ggtccgcagt
1141 tattgctatt tggctaccgt cagcgatctc tccaccaaag ctgcgtgccc gaccatggga
1201 gaagctcaca atgacaaacg tgctgaccca gcttttgtgt gcagacaagg agtggtggac
1261 aggggctggg gcaacggctg cggattattt ggcaaaggaa gcattgacac atgcgccaaa
1321 tttgcctgct ctaccaaggc aataggaaga accatcttga aagagaatat caagtacgaa
1381 gtggccattt ttgtccatgg accaactact gtggagtcgc acggaaacta ctccacacag
1441 gttggagcca ctcaggcagg gagattcagc atcactcctg cggcgccttc atacacacta
1501 aagcttggag aatatgagga ggtgacagtg gactgtgaac cacggtcagg gattgacacc
1561 aatgcatact acgtgatgac tgttggaaca aagacgttct tggtccatcg tgagtggttc
1621 atggacctca acctcccttg gagcagtgct ggaagtactg tgtggaggaa cagagagacg
1681 ttaatggagt ttgaggaacc acacgccacg aagcagtctg tgatagcatt gggctcacaa
1741 gagggagctc tgcatcaagc tttggctgga gccattcctg tggaattttc aagcaacact
1801 gtcaagttga cgtcgggtca tttgaagtgt agagtgaaga tggaaaaatt gcagttgaag
1861 ggaacaacct atggcgtctg ttcaaaggct ttcaagtttc ttgggactcc cgcagacaca
1921 ggtcacggca ctgtggtgtt ggaattgcag tacactggca cggatggacc ttgtaaagtt
1981 cctatctcgt cagtggcttc attgaacgac ctaacgccag tgggcagatt ggtcactgtc
2041 aacccttttg tttcagtggc cacggccaac gctaaggtcc tgattgaatt ggaaccaccc
2101 tttggagact catacatagt ggtgggcaga ggagaacaac agatcaatca ccattggcac
2161 aagtctggaa gcagcattgg caaagccttt acaaccaccc tcaaaggagc gcagagacta
2221 gccgctctag gagacacagc ttgggacttt ggatcagttg gaggggtgtt cacctcagtt
2281 gggaaggctg tccatcaagt gttcggagga gcattccgct tactgttcgg aggcatgtcc
2341 tggataacgc aaggattgct gggggctctc ctgttgtgga tgggcatcaa tgctcgtgat
2401 aggtccatag ctctcacgtt tctcgcagtt ggaggagttc tgctcttcct ctccgtgaac
2461 gtgcacgctc acactgggtg tgccatagac atcagccggc aagagctgag atgtggaagt
2521 ggagtgttca tacacaatga tgtggaggct tggatggacc gatacaagta ttaccctgaa
2581 acgccacaag gcctagccaa gatcattcag aaagctcata aggaaggagt gtgcggtcta
2641 cgatcagttt ccagactgga gcatcaaatg tgggaagcag tgaaggacga gctgaacact
2701 cttttgaagg agaatggtgt ggaccttagt gtcgtggttg agaaacagga gggaatgtac
2761 aagtcagcac ctaaacgcct caccgccacc acggaaaaat tggaaattgg ctggaaggcc
2821 tggggaaaga gtatttatt tgcaccagaa ctcgccaaca acacctttgt ggttgatggt
2881 ccggagacca aggaatgtcc gactcagaat cgcgcttgga atagcttaga agtggaggat
2941 tttggatttg gtctcaccag cactcggatg ttcctgaagg tcagagagag caacacaact
3001 gaatgtgact cgaagatcat tggaacggct gtcaagaaca acttggcgat ccacagtgac
3061 ctgtcctatt ggattgaaag caggctcaat gatacgtgga agcttgaaag ggcagttctg
3121 ggtgaagtca aatcatgtac gtggcctgag acgcatacct tgtggggcga tggaatcctt
3181 gagagtgact tgataatacc agtcacactg gcgggaccac gaagcaatca caatcggaga
3241 cctgggtaca agacacaaaa ccagggccca tgggacgaag gccgggtaga gattgacttc
3301 gattactgcc caggaactac ggtcaccctg agtgagagct gcggacaccg tggacctgcc
3361 actcgcacca ccacagagag cggaaagttg ataacagatt ggtgctgcag gagctgcacc
3421 ttaccaccac tgcgctacca aactgacagc ggctgttggt atggtatgga gatcagacca
3481 cagagacatg atgaaaagac cctcgtgcag tcacaagtga atgcttataa tgctgatatg
3541 attgaccctt ttcagttggg ccttctggtc gtgttcttgg ccacccagga ggtccttcgc
3601 aagaggtgga cagccaagat cagcatgcca gctatactga ttgctctgct agtcctggtg
3661 tttgggggca ttacttacac tgatgtgtta cgctatgtca tcttggtggg ggcagctttc
3721 gcagaatcta attcgggagg agacgtggta cacttggcgc tcatggcgac cttcaagata
3781 caaccagtgt ttatggtggc atcgtttctc aaagcgagat ggaccaacca ggagaacatt
3841 ttgttgatgt tggcggctgt tttctttcaa atggcttatc acgatgcccg ccaaattctg
3901 ctctgggaga tccctgatgt gttgaattca ctggcggtag cttggatgat actgagagcc
3961 ataacattca caacgacatc aaacgtggtt gttccgctgc tagccctgct aacacccggg
4021 ctgagatgct tgaatctgga tgtgtacagg atactgctgt tgatggtcgg aataggcagc
4081 ttgatcaggg agaagaggag tgcagctgca aaaaagaaag gagcaagtct gctatgcttg
4141 gctctagcct caacaggact tttcaacccc atgatccttg ctgctggact gattacatgt
4201 gatcccaacc gtaaacgcgg atggcccgca actgaagtga tgacagctgt cggcctgatg
4261 tttgccatcg tcggaggcgt ggcagagctt gacattgact ccatggccat tccaatgact
4321 atcgcggggc tcatgtttgc tgctttcgtg atttctggga aatcaacaga tatgtggatt
4381 gagagaacgg cggacatttc ctgggaaagt gatgcagaaa ttacaggctc gagcgaaaga
4441 gttgatgtgc ggcttgatga tgatggaaac ttccagctca tgaatgatcc aggagcacct
4501 tggaagatat ggatgctcag aatggtctgt ctcgcgatta gtgcgtacac cccctgggca
4561 atcttgccct cagtagttgg attttggata actctccaat acacaaagag aggaggcgtg
4621 ttgtgggaca ctccctcacc aaaggagtac aaaaaggggg acacgaccac cggcgtctac
4681 aggatcatga ctcgtgggct gctcggcagt tatcaagcag gagcgggcgt gatggttgaa
4741 ggtgtttttc acaccctttg gcatacaaca aaaggagccg ctttgatgag cggagagggc
4801 cgcctggacc catactgggg cagtgtcaag gaggatcgac tttgttacgg aggaccctgg
```

```
4861 aaattgcagc acaagtggaa cgggcaggat gaggtgcaga tgattgtggt ggaacctggc
4921 aagaacgtta agaacgtcca gacgaaacca ggggtgttca aaacacctga aggagaaatc
4981 ggggccgtga ctttggactt ccccactgga acatcaggct caccaatagt ggacaaaaac
5041 ggtgatgtga ttgggcttta tggcaatgga gtcataatgc ccaacggctc atacataagc
5101 gcgatagtgc agggtgaaag gatggatgag ccaatcccag ccggattcga acctgagatg
5161 ctgaggaaaa aacagatcac tgtactggat ctccatcccg gcgccggtaa aacaaggagg
5221 attctgccac agatcatcaa agaggccata aacagaagac tgagaacagc cgtgctagca
5281 ccaaccaggg ttgtggctgc tgagatggct gaagcactga gaggactgcc catccggtac
5341 cagacatccg cagtgcccag agaacataat ggaaatgaga ttgttgatgt catgtgtcat
5401 gctaccctca cccacaggct gatgtctcct cacagggtgc cgaactacaa cctgttcgtg
5461 atggatgagg ctcatttcac cgacccagct agcattgcag caagaggtta catttccaca
5521 aaggtcgagc taggggaggc ggcggcaata ttcatgacag ccaccccacc aggcacttca
5581 gatccattcc cagagtccaa ttcaccaatt tccgacttac agactgagat cccggatcga
5641 gcttggaact ctggatacga atggatcaca gaatacaccg ggaagacggt ttggtttgtg
5701 cctagtgtca agatggggaa tgagattgcc ctttgcctac aacgtgctgg aaagaaagta
5761 gtccaattga acagaaagtc gtacgagacg gagtacccaa aatgtaagaa cgatgattgg
5821 gactttgtta tcacaacaga catatctgaa atgggggcta actttaaggc gagcagggtg
5881 attgacagcc ggaagagtgt gaaaccaacc atcataacag aaggagaagg gagagtgatc
5941 ctgggagaac catctgcagt gacagcagct agtgccgccc agagacgtgg acgtatcggt
6001 agaaatccgt cgcaagttgg tgatgagtac tgttatgggg ggcacacgaa tgaagacgac
6061 tcgaacttcg cccattggac tgaggcacga atcatgctgg acaacatcaa catgccaaac
6121 ggactgatcg ctcaattcta ccaaccagag cgtgagaagg tatataccat ggatggggaa
6181 taccggctca gaggagaaga gagaaaaaac tttctggaac tgttgaggac tgcagatctg
6241 ccagtttggc tggcttacaa ggttgcagcg gctggagtgt cataccacga ccggaggtgg
6301 tgctttgatg gtcctaggac aaacacaatt ttagaagaca caacgaagt ggaagtcatc
6361 acgaagcttg gtgaaaggaa gattctgagg ccgcgctgga ttgacgccag ggtgtactcg
6421 gatcaccagg cactaaaggc gttcaaggac ttcgcctcgg gaaaacgttc tcagataggg
6481 ctcattgagg ttctgggaaa gatgcctgag cacttcatgg ggaagacatg ggaagcactt
6541 gacaccatgt acgttgtggc cactgcagag aaaggaggaa gagctcacag aatggccctg
6601 gaggaactgc cagatgctct tcagacaatt gccttgattg ccttattgag tgtgatgacc
6661 atgggagtat tcttcctcct catgcagcgg aagggcattg gaaagatagg tttgggaggc
6721 gctgtcttgg gagtcgcgac cttttttctgt tggatggctg aagttccagg aacgaagatc
6781 gccggaatgt tgctgctctc ccttctcttg atgattgtgc taattcctga gccagagaag
6841 caacgttcgc agacagacaa ccagctagcc gtgttcctga tttgtgtcat gaccccttgtg
6901 agcgcagtgg cagccaacga gatgggttgg ctagataaga ccaagagtga cataagcagt
6961 ttgtttgggc aaagaattga ggtcaaggag aatttcagca tgggagagtt tcttctggac
7021 ttgaggccgg caacagcctg gtcactgtac gctgtgacaa cagcggtcct cactccactg
7081 ctaaagcatt tgatcacgtc agattacatc aacacctcat tgacctcaat aaacgttcag
7141 gcaagtgcac tattcacact cgcgcgaggc ttccccttcg tcgatgttgg agtgtcggct
7201 ctcctgctag cagccggatg ctgggggacaa gtcaccctca ccgttacggt aacagcggca
7261 acactccttt tttgccacta tgcctacatg gttcccggtt ggcaagctga ggcaatgcgc
7321 tcagcccagc ggcggacagc ggccggaatc atgaagaacg ctgtagtgga tggcatcgtg
7381 gccacggacg tcccagaatt agagcgcacc acacccatca tgcagaagaa agttggacag
7441 atcatgctga tcttggtgtc tctagctgca gtagtagtga acccgtctgt gaagacagta
7501 cgagaagccg gaattttgat cacggccgca gcggtgacgc tttgggagaa tggagcaagc
7561 tctgtttgga acgcaacaac tgccatcgga ctctgccaca tcatgcgtgg gggttggttg
7621 tcatgtctat ccataacatg gacactcata aagaacatgg aaaaaccagg actaaaaaga
7681 ggtggggcaa aaggacgcac cttgggagag gtttggaaag aaagactcaa ccagatgaca
7741 aaagaagagt tcactaggta ccgcaaagag gccatcatcg aagtcgatcg ctcagcagca
7801 aaacacgcca ggaaagaagg caatgtcact ggagggcatc cagtctctag gggcacagca
7861 aaactgagat ggctggtcga acggaggttt ctcgaaccgg tcggaaaagt gattgacctt
7921 ggatgtggaa gaggcggttg gtgttactat atggcaaccc aaaaaagagt ccaagaagtc
7981 agagggtaca caaagggcgg tcccggacat gaagagcccc aactagtgca aagttatgga
8041 tggaacattg tcaccatgaa gagtgggggtg gatgtgttct acagaccttc tgagtgttgt
8101 gacaccctcc tttgtgacat cggagagtcc tcgtcaagtg ctgaggttga agagcatagg
8161 acgattcggg tccttgaaat ggttgaggac tggctgcacc gagggccaag ggaattttgc
8221 gtgaaggtgc tctgccccta catgccgaaa gtcatagaga agatggagct gctccaacgc
8281 cggtatgggg ggggactggt cagaaaccca ctctcacgga attccacgca cgagatgtat
8341 tgggtgagtc gagcttcagg caatgtggta cattcagtga atatgaccag ccaggtgctc
8401 ctaggaagaa tggaaaaaag gacctggaag ggaccccaat acgaggaaga tgtaaacttg
8461 ggaagtggaa ccagggcggt gggaaaaccc ctgctcaact cagacaccag taaaatcaag
8521 aacaggattg aacgactcag gcgtgagtac agttcgacgt ggcaccacga tgagaaccac
8581 ccatatagaa cctggaacta tcacggcagt tatgatgtga gcccacagg ctccgccagt
8641 tcgctggtca atggagtggt caggctcctc tcaaaaccat gggacaccat cacgaatgtt
8701 accaccatgg ccatgactga cactactccc ttcgggcagc agcgagtgtt caaagagaag
8761 gtggacacga aagctcctga accgccagaa ggagtgaagt acgtgctcaa cgagaccacc
```

```
 8821 aactggttgt gggcgttttt ggccagagaa aaacgtccca gaatgtgctc tcgagaggaa
 8881 ttcataagaa aggtcaacag caatgcagct ttgggtgcca tgtttgaaga gcagaatcaa
 8941 tggaggagcg ccagagaggc agttgaagat ccaaaatttt gggagatggt ggatgaggag
 9001 cgcgaggcac atctgcgggg ggaatgtcac acttgcattt acaacatgat gggaaagaga
 9061 gagaaaaaac ccggagagtt cggaaaggcc aagggaagca gagccatttg gttcatgtgg
 9121 ctcggagctc gctttctgga gttcgaggct ctgggttttc tcaatgaaga ccactggctt
 9181 ggaagaaaga actcaggagg aggtgtcgag ggcttgggcc tccaaaaact gggttacatc
 9241 ctgcgtgaag ttggcacccg gcctggggggc aagatctatg ctgatgacac agctggctgg
 9301 gacacccgca tcacgagagc tgacttggaa aatgaagcta aggtgcttga gctgcttgat
 9361 ggggaacatc ggcgtcttgc cagggccatc attgagctca cctatcgtca caaagttgtg
 9421 aaagtgatgc gcccggctgc tgatggaaga accgtcatgg atgttatctc cagagaagat
 9481 cagaggggga gtggacaagt tgtcacctac gccctaaaca ctttcaccaa cctggccgtc
 9541 cagctggtga ggatgatgga aggggaagga gtgattggcc cagatgatgt ggagaaactc
 9601 acaaaaggga aaggacccaa agtcaggacc tggctgtttg agaatgggga agaaagactc
 9661 agccgcatgg ctgtcagtgg agatgactgt gtggtaaagc ccctggacga tcgctttgcc
 9721 acctcgctcc acttcctcaa tgctatgtca aaggttcgca aagacatcca agagtggaaa
 9781 ccgtcaactg gatggtatga ttggcagcag gttccatttt gctcaaacca tttcactgaa
 9841 ttgatcatga aagatggaag aacactggtg gttccatgcc gaggacagga tgaattggta
 9901 ggcagagctc gcatatctcc aggggccgga tggaacgtcc gcgacactgc ttgtctggct
 9961 aagtcttatg cccagatgtg gctgcttctg tacttccaca gaagagacct gcggctcatg
10021 gccaacgcca tttgctccgc tgtccctgtg aattgggtcc ctaccggaag aaccacgtgg
10081 tccatccatg caggaggaga gtggatgaca acagaggaca tgttggaggt ctggaaccgt
10141 gtttggatag aggagaatga atggatggaa gacaaaaccc cagtggagaa atggagtgac
10201 gtcccatatt caggaaaacg agaggacatc tggtgtggca gcctgattgg cacaagagcc
10261 cgagccacgt gggcagaaaa catccaggtg gctatcaacc aagtcagagc aatcatcgga
10321 gatgagaagt atgtggatta catgagttca ctaaagagat atgaagacac aactttggtt
10381 gaggacacag tactgtagat atttaatcaa ttgtaaatag acaatataag tatgcataaa
10441 agtgtagttt tatagtagta tttagtggtg ttagtgtaaa tagtacggtg aattttgagg
10501 agaaagtcag gccgggaagt tcccgccacc ggaagttgag tagacggtgc tgcctgcgac
10561 tcaaccccag gaggactggg tgaacaaagc cgcgaagtga tccatgtaag ccctcagaac
10621 cgtctcggaa ggaggacccc acatgttgta acttcaaagc ccaatgtcag accacgctac
10681 ggcgtgctac tctgcggaga gtgcagtctg cgatagtgcc ccaggaggac tgggttaaca
10741 aaggcaaacc aacgccccac gcggccctag ccccggtaat ggcgttaacc agggcgaaag
10801 gactagaggt tagaggagac cccgcggttt aaagtgcacg gcccagcctg gctgaagctg
10861 taggtcaggg gaaggactag aggttagtgg agaccccgtg ccacaaaaca ccacaacaaa
10921 acagcatatt gacacctggg atagactagg agatcttctg ctctgcacaa ccagccacac
10981 ggcacagtgc gccgacaatg gtggctggtg gtgcgagaac acaggatct
```
//

Apr 17 2007 11:10:07

Search CoreNucleotide   for | Go | | Clear |

Limits   Preview/Index   History   Clipboard   Details

Display GenBank   Show 5   Send to   Hide: ☐ sequence   ☐ all but gene, CDS and mRNA features

Range: from begin   to end   ☐ Reverse complemented strand   Features: |+| | Refresh |

☐ **1:** AF404757. Reports West Nile virus i...[gi:21929240]   Links

Features   Sequence

```
LOCUS       AF404757               11029 bp ss-RNA     linear   VRL 23-JUL-2002
DEFINITION  West Nile virus isolate WN Italy 1998-equine, complete genome.
ACCESSION   AF404757
VERSION     AF404757.1  GI:21929240
KEYWORDS    .
SOURCE      West Nile virus
  ORGANISM  West Nile virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
            Flavivirus; Japanese encephalitis virus group.
REFERENCE   1  (bases 1 to 11029)
  AUTHORS   Lanciotti,R.S., Ebel,G.D., Deubel,V., Kerst,A.J., Murri,S.,
            Meyer,R., Bowen,M., McKinney,N., Morrill,W.E., Crabtree,M.B.,
            Kramer,L.D. and Roehrig,J.T.
  TITLE     Complete genome sequences and phylogenetic analysis of West Nile
            virus strains isolated from the United States, Europe, and the
            Middle East
  JOURNAL   Virology 298 (1), 96-105 (2002)
   PUBMED   12093177
REFERENCE   2  (bases 1 to 11029)
  AUTHORS   Deubel,V., Bowen,M., Meyer,R., McKinney,N. and Morrill,W.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Division of Vector-Borne Infectious
            Diseases, Centers for Disease Control & Prevention, Rampart Road,
            Fort Collins, CO 80521, USA
FEATURES             Location/Qualifiers
     source          1..11029
                     /organism="West Nile virus"
                     /mol_type="genomic RNA"
                     /isolate="WN Italy 1998-equine"
                     /specific_host="equine"
                     /db_xref="taxon:11082"
                     /country="Italy"
     CDS             97..10398
                     /note="contains capsid, pre-membrane, envelope, NS1, NS2a,
                     NS2b, NS3, NS4a, NS4b, and NS5"
                     /codon_start=1
                     /product="polyprotein precursor"
                     /protein_id="AAM81753.1"
                     /db_xref="GI:21929241"
                     /translation="MSKKPGGPGKSRAVNMLKRGMPRVLSLIGLKRAMLSLIDGKGPI
                     RFVLALLAFFRFTAIAPTRAVLDRWRGVNKQTAMKHLLSFKKELGTLTSAINRRSSKQ
                     KKRGGKTGIAVMIGLIASVGAVTLSNFQGKVMMTVNATDVTDVITIPTAAGKNLCIVR
                     AMDVGYMCDDTITYECPVLSAGNDPEDIDCWCTKSAVYVRYGRCTKTRHSRRSRRSLT
                     VQTHGESTLANKKGAWMDSTKATRYLVKTESWILRNPGYALVAAVIGWMLGSNTMQRV
                     VFVVLLLLVAPAYSFNCLGMSNRDFLEGVSGATWVDLVLEGDSCVTIMSKDKPTIDVK
                     MMNMEAANLAEVRSYCYLATVSDLSTKAACPTMGEAHNDKRADPAFVCRQGVVDRGWG
                     NGCGLFGKGSIDTCAKFACSTKATGRTILKENIKYEVAIFVHGPTTVESHGNYSTQIG
                     ATQAGRFSITPAAPSYTLKLGEYGEVTVDCEPRSGIDTNAYYVMTVGTKTFLVHREWF
                     MDLNLPWSSAGSTVWRNRETLMEFEEPHATKQSVIALGSQEGALHQALAGAIPVEFSS"
```

```
                    NTVKLTSGHLKCRVKMEKLQLKGTTYGVCSKAFKFLGTPADTGHGTVVLELQYTGTDG
                    PCKVPISSVASLNDLTPVGRLVTVNPFVSVATANAKVLIELEPPFGDSYIVVGRGEQQ
                    INHHWHKSGSSIGKAFTTTLKGAQRLAALGDTAWDFGSVGGVFTSVGKAVHQVFGGAF
                    RSLFGGMSWITQGLLGALLLWMGINARDRSIALTFLAVGGVLLFLSVNVHADTGCAID
                    ISRQELRCGSGVFIHNDVEAWMDRYKYYPETPQGLAKIIQKAHKEGVCGLRSVSRLEH
                    QMWESVKDELNTLLKENGVDLSVVVEKQEGMYKSAPKRLTATTEKLEIGWKAWGKSIL
                    FAPELANNTFVVDGPETKECPTQNRAWNSLEVEDFGFGLTSTRMFLKVRESNTTECDS
                    KIIGTAVKNNLAIHSDLSYWIESRLNDTWKLERAVLGEVKSCTWPETHTLWGDGILES
                    DLIIPVTLAGPRSNHNRRPGYKTQNQGPWDEGRVEIDFDYCPGTTVTLSESCGHRGPA
                    TRTTTESGKLITDWCCRSCTLPPLRYQTDSGCWYGMEIRPQRHDEKTLVQSQVNAYNA
                    DMIDPFQLGLLVVFLATQEVLRKRWTAKISMPAILIALLVLVFGGITYTDVLRYVILV
                    GAAFAESNSGGDVVHLALMATFKIQPVFMVASFLKARWTNQENILLMLAAVFFQMAYH
                    DARQILLWEIPDVLNSLAVAWMILRAITFTTTSNVVVPLLALLTPGLRCLNLDVYRIL
                    LLMVGIGSLIREKRSAAAKKKGASLLCLALASTGLFNPMILAAGLIACDPNRKRGWPA
                    TEVMTAVGLMFAIVGGLAELDIDSMAIPMTIAGLMFAAFVISGKSTDMWIERTADISW
                    ESDAEITGSSERVDVRLDDDGNFQLMNDPGAPWKIWMLRMACLAISAYTPWAILPSVI
                    GFWITLQYTKRGGVLWDTPSPKEYKKGDTTTGVYRIMTRGLLGSYQAGAGVMVEGVFH
                    TLWHTTKGAALMSGEGRLDPYWGSVKEDRLCYGGPWKLQHKWNGQDEVQMIVVEPGKN
                    VKNVQTKPGVFKTPEGEIGAVTLDFPTGTSGSPIVDKNGDVIGLYGNGVIMPNGSYIS
                    AIVQGERMDEPIPAGFEPEMLRKKQITVLDLHPGAGKTRRILPQIIKEAINRRLRTAV
                    LAPTRVVAAEMAEALRGLPIRYQTSAVTREHNGNEIVDVMCHATLTHRLMSPHRVPNY
                    NLFVMDEAHFTDPASIAARGYISTKVELGEAAAIFMTATPPGTSDPFPESNSPISDLQ
                    TEIPDRAWNSGYEWITEYIGKTVWFVPSVKMGNEIALCLQRAGKKVVQLNRKSYETEY
                    PKCKNDDWDFVITTDISEMGANFKASRVIDSRKSVKPTIITEGEGRVILGEPSAVTAA
                    SAAQRRGRIGRNPSQVGDEYCYGGHTNEDDSNFAHWTEARIMLDNINMPNGLIAQFYQ
                    PEREKVYTMDGEYRLRGEERKNFLELLRTADLPVWLAYKVAAAGVSYHDRRWCFDGPR
                    TNTILEDNNEVEVITKLGERKILRPRWIDARVYSDHQALKAFKDFASGKRSQIGLIEV
                    LGKMPEHFMGKTWEALDTMYVVATAEKGGRAHRMALEELPDALQTIALIALLSVMTMG
                    VFFLLMQRKGIGKIGLGGVVLGVATFFCWMAEVPGTKIAGMLLLSLLLMIVLIPEPEK
                    QRSQTDNQLAVFLICVMTLVSAVAANEMGWLDKTKSDISSLFGQRIEVKENFSMGEFL
                    LDLRPATAWSLYAVTTAVLTPLLKHLITSDYINTSLTSINVQASALFTLARGFPFVDV
                    GVSALLLAAGCWGQVTLTVTVTAATLLFCHYAYMVPGWQAEAMRSAQRRTAAGIMKNA
                    VVDGIVATDVPELERTTPIMQKKVGQIMLILVSLAAVVVNPSVKTVREAGILITAAAV
                    TLWENGASSVWNATTAIGLCHIMRGGWLSCLSITWTLIKNMDKPGLKRGGAKGRTLGE
                    VWKERLNQMTKEEFTRYRKEAIIEVDRSAAKHARKEGNVTGGHPVSRGTAKLRWLVER
                    RFLEPVGKVIDLGCGRGGWCYYMATQKRVQEVRGYTKGGPGHEEPQLVQSYGWNIVTM
                    KSGVDVFYRPSECCDTLLCDIGESSSSAEVEEHRTIRVLEMVEDWLHRGPREFCVKVL
                    CPYMPKVIEKMELLQRRYGGGLVRNPLSRNSTHEMYWVSRASGNVVHSVNMTSQVLLG
                    RMEKRTWKGPQYEEDVNLGSGTRAVGKPLLNSDTSKIKNRIERLRREYSSTWHHDENH
                    PYRTWNYHGSYDVKPTGSASSLVNGVVRLLSKPWDTITNVTTMAMTDTTPFGQQRVFK
                    EKVDTKAPEPPEGVKYVLNETTNWLWAFLAREKRPRMCSREEFIRKVNSNAALGAMFE
                    EQNQWRSAREAVEDPKFWEMVDEEREAHLRGECHTCIYNMMGKREKKPGEFGKAKGSR
                    AIWFMWLGARFLEFEALGFLNEDHWLGRKNSGGGVEGLGLQKLGYILREVGTRPGGKI
                    YADDTAGWDTRITRADLENEAKVLELLDGEHRRLARAIIELTYRHKVVKVMRPAADGR
                    TVMDVISREDQRGSGQVVTYALNTFTNLAVQLVRMMEGEGVIGPDDVEKLTKGKGPKV
                    RTWLFENGEERLSRMAVSGDDCVVKPLDDRFATSLHFLNAMSKVRKDIQEWKPSTGWY
                    DWQQVPFCSNHFTELIMKDGRTLVVPCRGQDELVGRARISPGAGWNVRDTACLAKSYA
                    QMWLLLYFHRRDLRLMANAICSAVPVNWVPTGRTTWSIHAGGEWMTTEDMLEVWNRVW
                    IEENEWMEDKTPVEKWSDVPYSGKREDIWCGSLIGTRARATWAENIQVAINQVRAIIG
                    DEKYVDYMSSLKRYEDTTLVEDTVL"
ORIGIN
        1 agtagttcgc ctgtgtgagc tgacaaactt agtagtgttt gtgaggatta acaacaatta
       61 acacagtgcg agctgtttct tagcacgaag atctcgatgt ctaagaaacc aggagggccc
      121 ggcaagagcc gggctgtcaa tatgctaaaa cgcggaatgc cccgcgtgtt gtccttgatt
      181 ggactgaaga gggcaatgtt gagcctgatc gacggcaagg ggccaatacg atttgtgttg
      241 gctctcttgg cgttcttcag gttcacagca attgctccga cccgagcagt gctggatcga
      301 tggagaggtg tgaacaaaca aacagcgatg aaacaccttc tgagtttttaa gaaagaacta
      361 gggaccttga ccagtgctat caatcggcgg agctcaaaac aaaagaaaag aggaggaaag
      421 accggaattg cagtcatgat tggcctgatc gccagcgtag gagcagttac cctctctaac
      481 tttcaaggga aagttatgat gacggtaaat gctactgacg tcacagatgt catcacgatt
      541 ccaacagctg ctggaaagaa tctatgtatt gtcagagcaa tggatgtggg atacatgtgt
      601 gatgacacta tcacttatga atgcccagtg ctgtcggccg gtaatgatcc agaagacatc
      661 gactgttggt gcacaaagtc agcagtctac gtcaggtatg gaagatgcac caagacacgc
      721 cactcaagac gcagtcggag gtcactgaca gtgcagacac acggagaaag cactctagcg
      781 aacaaaaagg gggcttggat ggacagcacc aaggccacaa ggtatctggt gaaaacagaa
      841 tcatggatcc tgaggaaccc tggatatgcc ctggtggcag ctgtcattgg ttggatgctt
```

```
 901 ggaagcaaca ccatgcagcg agttgtgttt gtcgtgttat tgctcctggt ggctccagcc
 961 tacagcttca actgccttgg aatgagcaac agagacttct tggagggagt gtctggagca
1021 acatgggtgg atttggttct cgaaggcgac agctgcgtga ctatcatgtc caaggacaag
1081 cccaccattg atgtgaagat gatgaatatg gaggctgcca acctggcaga ggtccgcagt
1141 tattgctatt tggctaccgt cagcgatctc tccaccaaag ctgcatgccc gaccatggga
1201 gaagctcaca atgacaaacg tgctgaccca gcttttgtgt gcagacaagg agtggtggac
1261 aggggctggg gcaacggctg cggactattt ggcaaaggaa gcattgacac atgcgccaaa
1321 tttgcctgct ccaccaaggc aacaggaaga accatcttga aagagaatat caagtatgaa
1381 gtggccatct ttgtccatgg accaaccact gtggagtcgc atggaaacta ctccacacag
1441 attggggcca ctcaggcagg gagattcagc atcactcctg cggcgccttc atacacacta
1501 aagcttggag aatatggaga agtgacagtg gactgtgaac cacggtcagg gattgacact
1561 aatgccctact acgtgatgac tgttggaaca aagacgtttt tggtccatcg tgagtggttc
1621 atggacctca acctcccttg gagcagtgct ggaagtactg tgtggaggaa cagagagacg
1681 ttgatggagt ttgaggaacc acacgccaca aagcagtctg tgatagcatt gggctcacaa
1741 gagggagctc tgcatcaagc tttggctgga gccatccctg tggaatttttc aagcaacact
1801 gtcaagttga cgtcgggcca tttgaagtgt agagtgaaga tggaaaaatt gcagttgaag
1861 ggaacaactt acggcgtctg ttcaaaggct ttcaagtttc ttgggactcc cgcagacaca
1921 ggtcacggca ctgtggtgtt ggaattgcag tacactggca cggatggacc ttgcaaagtt
1981 cccatctcgt cagtggcttc attgaacgac ctaacaccgg taggcagatt ggtcactgtt
2041 aacccttttg tttcagtggc cacggccaac gccaaggtcc tgattgaact ggaaccaccc
2101 tttggagact catacatagt ggtaggcaga ggagaacaac agatcaatca ccattggcat
2161 aagtctggaa gcagcattgg caaagccttt acaaccactc tcaaaggggc gcagagatta
2221 gccgctctag gagacacagc ttgggacttc ggatcagttg gaggggtgtt tacctcagta
2281 gggaaggctg tccatcaagt gttcggtgga gcattccgct cactgttcgg aggtatgtcc
2341 tggataacgc agggattgct gggggctctt ctgttgtgga tgggcatcaa tgctcgtgac
2401 aggtccatag ctctcacgtt tctcgcagtt ggaggagtcc tgctcttcct ctccgtgaac
2461 gtgcacgctg acactgggtg tgccatagac atcagccggc aagagctgag atgtggaagt
2521 ggagtgttca tacacaatga tgtggaggct tggatggacc gatacaagta ctaccctgaa
2581 acgccacaag gcctagccaa gattattcaa aaagcccata aggaaggagt gtgcggtcta
2641 cggtcagttt ccagactgga gcatcaaatg tgggaatcag tgaaggacga gctgaacact
2701 cttttgaaag agaatggtgt ggaccttagt gtcgtggttg agaaacagga gggaatgtac
2761 aagtcggcac ctaaacgtct caccgccacc acggaaaaat tggaaattgg ctggaaggcc
2821 tggggaaaga gcattctatt tgcaccagaa cttgccaaca acacctttgt ggttgatggt
2881 ccggagacca aggaatgccc aactcagaat cgcgcttgga atagtttgga agtggaggat
2941 tttgggtttg gtctcaccag cactcggatg ttcctgaagg tcagagagag caacacaact
3001 gaatgtgact cgaagatcat cggaacggcc gtcaaaaaca acttggcgat ccacagtgac
3061 ctgtcctatt ggattgaaag caggctcaat gatacgtgga agcttgaaag ggcggttctg
3121 ggtgaagtca aatcatgcac gtggcctgag acgcacacct tgtggggcga tggaatcctc
3181 gagagtgacc tgataaatacc agtcacactg gcgggaccac gaagcaacca caatcggaga
3241 cctgggtaca agacacaaaa ccagggccca tgggacgaag gtcgagtaga gattgatttc
3301 gattactgcc cagggaccac ggtcaccctg agtgagagct gcggacaccg tggacctgcc
3361 actcgcacca ccacagagag cggaaagttg ataacagatt ggtgctgcag gagctgcacc
3421 ttaccaccat tgcgctacca gacggacagc ggctgttggt atggcatgga gattagacca
3481 cagagacatg atgaaaagac tcttgtgcag tcacaagtga atgcttacaa tgctgatatg
3541 attgatcctt ttcagctggg ccttctggtc gtgttcttgg ccacccagga ggtccttcgc
3601 aagaggtgga cagccaagat cagcatgcca gctatactga ttgctctgct agtcctggtg
3661 tttgggggca ttacctacac tgatgtgttg cgctatgtca tcttggtggg agcagctttc
3721 gcagaatcca attcaggagg agacgtggta cacttggcgc tcatggcgac cttttaagata
3781 cagccagtgt ttatggtggc atcatttctc aaagcgagat ggaccaacca ggagaacatc
3841 ttgttgatgt tggcggctgt tttcttccaa atggcttacc acgatgcccg ccaaattctg
3901 ctctgggaga tccctgatgt gttgaactca ctggcagtag cttggatgat actgagagcc
3961 ataacattca caacgacatc aaacgtggtt gttccgttgc tggcccttct aacacccggg
4021 ctgagatgct tgaatctgga tgtgtacagg attctgctgt tgatggtcgg aataggcagt
4081 ttgatcaggg agaagagaag tgcagctgca aaaaagaaag gagcaagtct gttatgcttg
4141 gctctagcct caacaggact tttcaacccc atgatccttg ccgctggact gattgcatgt
4201 gatcccaacc gcaaacgcgg gtggccccgca actgaagtga tgactgctgt cggcttgatg
4261 tttgccattg tcggagggct ggcagagctt gacattgact ccatggccat tccaatgact
4321 atcgcggggc tcatgtttgc tgctttcgtg atttctggga atcaacaga tatgtggatc
4381 gagaggacgg cggacatctc ctgggaaagt gatgcagaaa tcacaggctc gagcgaaaga
4441 gttgatgtgc ggcttgatga tgatggaaac ttccagctca tgaatgatcc aggagcacct
4501 tggaaaatat ggatgctcag aatggcctgc ctcgcgatta gtgcgtacac cccttgggcg
4561 atcctgccct cagtaatcgg attttggata actctccaat acacaaagag aggaggcgtg
4621 ctgtgggaca ctccctcacc aaaggagtac aaaaaggggg acacgaccac cggagtctac
4681 aggattatga ctcgtgggct gctcggcagt tatcaagcag gagcgggcgt gatggttgaa
4741 ggcgtttcc acactctttg gcacacaaca aaaggagccg ctttgatgag cggagaaggc
4801 cgcctggacc catactgggg tagtgtcaag gaggaccgac tttgttacgg aggaccctgg
```

```
4861 aaattgcagc ataagtggaa cgggcaggat gaggtgcaga tgattgtggt agaacccggc
4921 aagaacgtca aaaacgtcca gacgaaacca ggggtgttca aaacacctga aggagaaatt
4981 ggggccgtga ctttggactt ccccactgga acatcaggct cacccatagt ggacaaaaac
5041 ggtgatgtga tcgggctcta tggcaatgga gtcataatgc ccaacggctc ctacataagc
5101 gcgatagtgc agggtgaaag gatggatgag ccaatcccag ccggattcga acctgagatg
5161 ctgaggaaaa aacagatcac tgttctggat cttcatcccg gtgccggtaa aacaaggagg
5221 attctgccac agatcattaa agaggccata aatagaagac tgagaacagc cgtgctagca
5281 ccaactaggg ttgtggctgc tgagatggct gaagccctga gaggactgcc tatccggtat
5341 cagacatctg cagtgaccag agaacacaat ggaaatgaga ttgttgatgt catgtgccat
5401 gccaccctca ctcacaggct gatgtctcct cacagggtgc ctaactacaa tctgtttgtg
5461 atggatgagg ctcacttcac cgatccagct agcatcgcag caagaggtta tatttccaca
5521 aaggtcgacg tggggaggc ggcggcaata ttcatgacag ccactccgcc gggcacctca
5581 gatccattcc cagagtccaa ttcacctatc tctgacttac agactgagat cccggatcgg
5641 gcctggaact ctggatacga atggatcaca gaatacattg ggaagacggt ttggttcgtg
5701 cctagtgtga aaatggggaa tgagattgcc ctttgtctac aacgtgccgg aaagaaagta
5761 gtccaattga acagaaagtc gtatgagacg gagtacccaa aatgtaagaa cgatgattgg
5821 gactttgtta tcacaacaga catatctgaa atgggggcta acttcaaggc gagcagggtg
5881 atcgacagca ggaagagcgt gaagccaacc atcataacag aaggagaagg gagagtgatc
5941 ctgggagaac catctgcagt gacggcagct agtgccgccc agagacgtgg acgcatcggt
6001 agaaatccgt cgcaagttgg tgatgagtac tgttatggag ggcacacgaa tgaagacgac
6061 tcgaacttcg cccattggac tgaggcacga atcatgctgg acaacatcaa catgccaaac
6121 ggactgatcg ctcaattcta tcaaccagag cgtgagaagg tatacaccat ggatggggaa
6181 taccggctca gaggagaaga gaggaaaaac ttcctggaat tgttgaggac tgcagatctg
6241 ccagtttggc tggcttataa ggtggcagcg gctggagtgt catacacga tcggaggtgg
6301 tgttttgatg gccctaggac aaatacgatc ttagaagaca acaacgaagt ggaagtcatc
6361 acgaagcttg gtgaaaggaa gattctgagg ccccgctgga ttgacgccag ggtgtactcg
6421 gatcaccagg cactgaaggc gttcaaggac tttgcttcgg ggaagcgttc tcagataggg
6481 ctcattgagg tcctgggaaa gatgcctgag catttcatgg ggaagacatg ggaagcactt
6541 gacaccatgt acgttgtggc cactgcagag aaaggaggaa gagctcacag aatggccctg
6601 gaggaactgc cggatgctct ccagacaatt gccttgattg ccttactgag tgtgatgacc
6661 atgggagtat tctttctcct catgcagcgg aagggcattg gaaagatagg tctgggaggc
6721 gttgtcttgg gagtcgcgac cttcttctgt tggatggctg aagttcccgg aacgaagatc
6781 gccggaatgt tgctgctttc ccttctcttg atgattgtgc taattcctga gccagagaag
6841 caacgttcgc agacagacaa ccagctagcc gtgttcctga tttgtgtcat gacccttgtg
6901 agcgcagtgg cagccaacga aatgggttgg ctagataaga ccaagagtga cataagcagt
6961 ttgtttgggc aaagaattga ggtcaaggag aatttcagca tgggagagtt tcttctggac
7021 ttgagaccgg caacagcctg gtcactgtac gctgtgacca cagcggtcct cactccactg
7081 ctaaagcatc tgatcacgtc agattacatt aacacttcat tgacctcaat taatgttcaa
7141 gcaagtgcac tattcacact cgcgcgaggc ttcccctttg tcgatgttgg agtgtcagct
7201 ctcttgctag cagccggatg ctgggacaa gtcaccctca ccgtgacggt gacagcggca
7261 acacttctgt tctgccacta tgcctacatg gttccggat ggcaggctga ggcaatgcgc
7321 tcagcccagc ggcggacagc ggctggaatc atgaaaaacg ctgtagtgga tggcatcgtg
7381 gccacggacg tcccagaatt agagcgcacc acacccatca tgcagaagaa agtcggacaa
7441 atcatgctga tcttggtgtc tctagctgca gtagtagtga accgtctgt gaagacagtg
7501 cgagaagctg gaattttgat cacggcagca gcggtgacgc tttgggagaa tggagcaagc
7561 tctgtttgga acgcaacaac tgccattgga ctctgccaca tcatgcgtgg gggttggttg
7621 tcatgcctat ccataacatg gacactcata aagaacatgg acaaaccagg actaaaaaga
7681 ggtggggcaa aaggacgcac cttgggagag gtttggaagg aaagactcaa ccagatgaca
7741 aaagaggagt tcactaggta ccgcaaagag gccatcatcg aagtcgatcg ctcagcagca
7801 aaacacgcca ggaaagaagg caatgtcacc ggagggcatc cagtctctag aggcacagca
7861 aaactgagat ggctggtcga acggaggttt ctcgaaccgg tcggaaaagt gattgacctt
7921 ggatgtggaa gaggcggttg gtgttactat atggcaaccc aaaaaagagt tcaagaagtc
7981 agagggtaca caaagggcgg tcccggacat gaagagcccc aactagtgca aagttatgga
8041 tggaacattg tcaccatgaa gagtggagtg gatgtgttct acagaccttc tgagtgttgt
8101 gacactctcc tttgtgacat cggagagtct tcgtcaagtg ctgaggttga agagcatagg
8161 acgattcggg tccttgaaat ggttgaggac tggctgcacc gagggccaag ggaattttgt
8221 gtgaaggtgc tctgcccta catgccaaaa gtcatagaga agatggagtt actccagcgc
8281 cggtatgggg gaggactggt cagaaaccca ctctcacgga attccacgca cgagatgtac
8341 tgggtaagtc gagcttcagg caatgtagta cattcagtga acatgactag ccaggtgctt
8401 ctgggaagga tggagaagag gacctggaag ggaccccaat acgaggaaga tgtaaacttg
8461 ggaagtggaa ccaggggcgt ggggaaaccc ctactcaact cagacaccag caaaatcaag
8521 aacaggattg aacgactcag gcgtgagtac agttcaacgt ggcaccacga tgagaaccac
8581 ccatatagaa cctgaactat cacggcagt tatgacgtga agcccacagg ctccgctagc
8641 tcgctggtca tggagtggt taggctcctc tcaaaaccat gggacaccat cacgaatgtc
8701 accaccatgg ccatgactga cacaactccc ttcggacagc agcgggtgtt caaagagaag
8761 gtggacacga aagctcctga accgccagaa ggagtgaagt atgtgctcaa tgaaaccacc
```

```
 8821 aattggttgt gggcgtttct ggccagagaa aaacgtccca gaatgtgctc tcgagaggaa
 8881 ttcataagaa aggtcaacag caatgcagct ttgggtgcca tgtttgaaga acagaatcaa
 8941 tggaggagcg ccagagaagc agttgaagat ccaaaatttt gggagatggt agatgaggag
 9001 cgcgaggcac acctgcgggg ggaatgtcac acttgcatct acaacatgat gggaaagaga
 9061 gagaaaaaac ccggagagtt cggaaaggcc aagggaagca gagccatatg gttcatgtgg
 9121 ctcggagctc gctttctgga gttcgaagct ctgggctttc tcaatgaaga tcactggctt
 9181 ggaagaaaga actcaggagg aggtgtcgag ggcttgggcc tccaaaaact gggttacatc
 9241 ctgcgtgaag ttggcacccg acctgggggc aagatctatg ctgacgacac agctggctgg
 9301 gacacccgca tcacgagagc tgacttggaa aatgaagcta aggtgcttga gctgcttgat
 9361 ggggaacatc ggcgtcttgc cagggccatc attgagctca cctatcgtca caaagttgtg
 9421 aaagtgatgc gcccggctgc tgatggaaga actgtcatgg atgttatctc cagagaagat
 9481 cagagaggga gcggacaagt tgtcacctac gccctaaaca ctttcaccaa cctggccgtc
 9541 cagctggtga ggatgatgga aggggaagga gtgattggcc cagatgatgt ggagaaactc
 9601 acaaagggga aaggacccaa agtcaggacc tggctgtttg agaatgggga agagagactc
 9661 agccgcatgg ctgtcagcgg agatgactgc gtggtaaagc ctctggatga tcgctttgcc
 9721 acttcgctcc actttctcaa cgccatgtca aaggtccgca aagacattca agagtggaaa
 9781 ccatcaactg gatggtatga ttggcagcag gttccattct gctcaaacca tttcactgaa
 9841 ttgatcatga aagatggaag aacattggtg gttccatgcc gaggacagga cgaactggta
 9901 ggcagggctc gcatttctcc aggggccgga tggaacgtcc gcgacaccgc ttgtctggct
 9961 aagtcatacg cccagatgtg gctgcttctg tacttccaca gaagagacct gcggctcatg
10021 gccaacgcca tttgctccgc tgtccctgtg aactgggtcc ccaccggaag aaccacatgg
10081 tctatccacg caggaggaga gtggatgaca acagaggaca tgctggaggt ttggaaccgt
10141 gtctggatag aggagaatga atggatggaa gacaaaaccc cagtggagaa atggagtgac
10201 gtcccatact caggaaaacg ggaggacatc tggtgtggca gcctgattgg cacaagagcc
10261 cgagccacgt gggcagaaaa catccaggtg gctatcaacc aagtcagagc aatcattgga
10321 gatgagaagt atgtggatta tatgagttca ctaaagagat atgaagacac aactttggtt
10381 gaggacacag tactgtaaac actttatcaa ttgtaaatag acaacataag catgtatata
10441 ggtataggtt tatagtggta tttaatgatg ttagtgtaga tagttaagaa aattttaagg
10501 aggaagtcag gccggaaaat ttccgccacc ggaagttgag tagacggtgc tgcctgcgac
10561 tcaaccccag gaggactggg tgaacaaagc cgtgaagtga tccatgtaag ccctcagaac
10621 cgtctcggaa ggaggacccc acatgttgta acttcaaagc ccaatgtcag accacgctac
10681 ggcgtgccac tctgcggaga gtgcagtctg cgatagtgcc ccaggaggac tgggttaaca
10741 aaggcaaacc aacgccccac gcggccctag ccctggtaat ggtgttaacc agggcgaaag
10801 gactagaggt tagaggagac cccgcggttt gaagagcacg gcccagcctg gctgaagctg
10861 taggtcaggg gaaggactag aggttagtgg agaccccgtg ccacaaaaca ccacaacaaa
10921 acagcatatt gacacctggg atagactagg agatcttctg ctctgcacaa ccagccacac
10981 ggcacagtgc gccgacaatg gtggctggtg gtgcgagaac acaggatct
```

//

Apr 17 2007 11:10:07